Genetic tools for high yield protein production with *Bacillus megaterium*

Von der Fakultät für Lebenswissenschaften

der Technischen Universität Carolo-Wilhelmina

zu Braunschweig

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

genehmigte

D i s s e r t a t i o n

von Simon Stammen

aus Kevelaer

1. Referent: Professor Dr. Dieter Jahn
2. Referent: Professor Dr. Petra Dersch
eingereicht am: 01.09.2010
mündliche Prüfung (Disputation) am: 08.10.2010

Druckjahr 2010

Vorveröffentlichungen der Dissertation:

Teilergebnisse aus dieser Arbeit wurden mit Genehmigung der Fakultät für Lebenswissenschaften, vertreten durch den Mentor der Arbeit, in folgenden Beiträgen vorab veröffentlicht:

Publikationen:

Stammen, S., Schuller, F., Dietrich, S., Gamer, M., Biedendieck R. and D. Jahn. 2010.
Application of *Escherichia coli* phage K1E DNA-dependent RNA polymerase for *in vitro* RNA synthesis and *in vivo* protein production in *Bacillus megaterium*.
Appl Microbiol Biotechnol **88**:529-539.

Stammen, S., Müller, B. K., Korneli, C., Biedendieck, R., Gamer, M., Franco-Lara, E. and D. Jahn. 2010.
High yield intra- and extracellular protein production using *Bacillus megaterium*.
Appl Environ Microbiol **76**:4037-4046.

Bunk, B., Schulz, A., Stammen, S., Münch, R., Warren, M. J., Rohde, M., Jahn, D. and R. Biedendieck. 2010.
A short story about a big magic bug.
Bioengineered Bugs **1**:1-7.

Gamer, M., Fröde, D., Biedendieck, R., Stammen, S. and D. Jahn. 2009.
A T7 RNA polymerase-dependent gene expression system for *Bacillus megaterium*.
Appl Microbiol Biotechnol **82**:1195-1203.

Tagungs- und Industriebeiträge:

Stammen, S., Schuller, F., Dietrich, S., Gamer, M., Biedendieck, R. and D. Jahn. 2010.
A novel DNA-dependent RNA polymerase of *Escherichia coli* phage K1E: production, characterization and application. (Poster)
World Congress on Industrial Biotechnology & Bioprocessing, Washington D.C., DC, USA.

Stammen, S., Müller, B. K., Hellert, J., Busch, M., Biedendieck, R. and D. Jahn. 2010.
Construction of a high performance plasmid for intra- and extracellular protein production in *Bacillus megaterium*. (Vortrag)
3. Gemeinsame Tagung der Vereinigung für allgemeine und angewandte Mikrobiologie (VAAM) und der deutschen Gesellschaft für Hygiene und Mikrobiologie (DGHM), Hannover, Deutschland.

S. Stammen. 2010.
Advances in heterologous protein production in *Bacillus megaterium*. (Vortrag)
Technologietransfergespräche beim Kooperationspartner Codexis Inc., Redwood City, CA, USA.

Stammen, S., Gamer, M., Müller, B. K. and D. Jahn. 2009.
A Plasmid borne system for high level protein production in *Bacillus megaterium*. (Poster)
Symbiosis, the 14th European Congress on Biotechnology, Barcelona, Spain.

Stammen, S., Gamer, M., Müller, B. K., Schwerk, J., Schlums, H. and D. Jahn. 2009.
Improved plasmids for heterologous protein production in *Bacillus megaterium*. (Poster)
Jahrestagung der Vereinigung für allgemeine und angewandte Mikrobiologie (VAAM), Bochum, Deutschland.

Stammen, S., Gamer, M., Müller, B. K., Schwerk, J., Schlums, H. and D. Jahn. 2008.
Improved plasmids for heterologous protein production in *Bacillus megaterium*. (Poster)
6th Conference on Recombinant Protein Production, Alghero, Sardina, Italy.

Stammen, S., Gamer, M., Schwerk, J., Schlums, H. and D. Jahn. 2008.
Construction of strong inducible promoters for *Bacillus megaterium* expression vectors. (Poster)
Jahrestagung der Vereinigung für allgemeine und angewandte Mikrobiologie (VAAM), Frankfurt, Deutschland.

Gamer, M., Stammen, S., Biedendieck, R., Yang, Y., Fröde, D., Bendig, C. and D. Jahn. 2007.
Bacillus megaterium - an alternative expression system. (Poster)
Symbiosis, the 13th European Congress on Biotechnology, Barcelona, Spain.

In den Wissenschaften ist viel Gewisses,
sobald man sich von den Ausnahmen nicht irremachen lässt
und die Probleme zu ehren weiß.

(Wilhelm Meisters Wanderjahre - Buch 2, J. W. von Goethe, 1749-1832)

Contents

Abbreviations

ASP	computationally designed artificial signal peptide
ATP	adenosine 5'-triphosphate
BSA	bovine serum albumine
bp	base pairs
CAI	codon adaptation index
CDW	cell dry weight
dH_2O	destilled water
ΔA	change in absorbance
d	thickness
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
dsDNA	double stranded DNA
DSMZ	German Collection of Microorganisms and Cell Cultures
DTT	1,4-dithio-D,L-threitol
EA	volumetric enzyme activity
EDTA	ethylenediaminetetraacetic acid
e.g.	*exempli gratia* (for example)
ε	extinction coefficient of p-nitrophenol (9.62 cm^2 μmol^{-1})
et al.	*et alteri* (and others)
for	forward
g	*centrifugation*: earth gravity
g	*weight*: gram
Gfp	green fluorescent protein
h	hour
His_6	polyhistidine tag
hp	high performance
IPTG	isopropyl-β-D-thiogalactopyranoside
kb	kilo base pairs
L	liter
LB	Luria Bertani
LPS	lipopolysaccharide
m	milli
M	molar [mol L^{-1}]
µ	micro

Mbp	mega base pairs
mcs	multiple cloning site
min	minute
MOPS	3-(N-morpholino)-propan sulfonacid
M_r	relative molecular weight
n	nano
OD_λ	optical density at wavelength λ
orf	open reading frame
ori	origin of replication
PCR	polymerase chain reaction
PEG 6000	polyethylen glycol with M_r of 6,000
PET	polyethylene terephtalate
*p*NPP	*p*-nitrophenylpalmitate
rbs	ribosome binding site
rev	reverse
RNA	ribonucleic acid
RNase	ribonuclease
RNAP	RNA polymerase
rpm	revolutions per minute
RT	room temperature
s	second
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SP	signal peptide
TEMED	N,N,N',N',-tetramethylen diamine
TEV	tobacco etch virus
T_m	melting temperature
Tris	tris-(hydroxymethyl)-aminomethane
U	units
utr	5' untranslated mRNA region
UV	ultraviolet
V_E	reaction volume
V_R	volume of enzyme solution
v/v	volume per volume
w/v	weight per volume

I. Introduction

I.1 Biotechnology

Microorganisms like lactic acid bacteria and yeasts were applied in fermentations of food and beverages (sourdough, beer or wine) centuries before the scientific background of these processes was understood. Since many decades, an ever-growing repertoire of modern molecular techniques allow for a steady growth in knowledge of chemical and biological principles employed by nature. Meanwhile, not only cultivated natural microbes and their products are used in biotechnological applications, rather more and more genetically modified organisms generate products with improved features. Next to high-priced pharmaceutical agents, an increasing amount of consumer products containing ingredients biotechnologically made are part of our modern life. Many of these products like packages made of bio-degradable biopolymers and various preservatives are intended to increase product sustainability. The use of enzymes in a wide range of daily applications (*e.g.* as additives in detergents or in food processing) reduces the energy consumption of the corresponding processes. Enzymes are highly selective biocatalysts and often even small quantities of these proteins are sufficient to catalyze chemical reactions at ambient temperatures and mild reaction conditions. Due to their chiral catalysis, enzymes are the alternative to complex chemical syntheses of regio- and stereoselective compounds, which are needed in the pharmaceutical and fine-chemical industry.

According to the DECHEMA, industrial biotechnological processes actually generate goods worth 55 billion Euros per year (Beller *et al.*, 2010). In 2004, its product sales were already five percent of the whole turnover of the chemical industry in accordance to a trend survey conducted by McKinsey & Company (Riese & Bachmann, 2004).

Representatives of the world's leading chemical companies, analysts as well as politicians are convinced that the industrial biotechnology will further impact industrial production processes. The oil spill in the Gulf of Mexico this year showed dramatically the necessity for alternatives to the production of natural oil, which in the future will only be possible from the few remaining, barely accessible sources. Thus,

the production of biofuels and biomaterials aiming at the reduction of environmental pollution and resource consumption by building up a sustainable biomass-based value chain is strongly intended. Therefore, the long-term goal of industrial biotechnology is the competitive generation of energy from renewable resources, replacement of the oil-based production pipelines by bioprocesses and the creation of high-value bioproducts with unique characteristics (Atiyeh, 2010; Keeling, 2010).

I.2 Recombinant protein production

To enable the production of desired proteins and enzymes for biotechnological applications nowadays often genetic techniques are employed. The target protein's gene is cloned into an expression host system which enables the production of sufficient amounts of functional product. To fulfill the quality and quantity criteria for an economic process an appropriate host system has to be chosen. If for example a pharmaceutical protein with modifications (*e.g.* glycolizations) similar to those of human proteins is desired, mammalian cell cultures are often the host of choice. However, this applies just for a few high-priced pharmaceutical proteins because the cultivations of higher eukaryotic cells are very cost intensive and often difficult to handle. Cheaper eukaryotic host alternatives are the baculovirus expression system with insect cells or yeast based systems like *Pichia pastoris* or *Saccharomyces cerevisiae*. Further, genetic tools for heterologous protein production using filamentous fungi were developed. In this context, a plasmid toolbox for gene expression in *Aspergillus niger* was established by members of the "Sonderforschungsbereich 578" at the Technische Universität Braunschweig (Roth & Dersch, 2010). Even though there are diverse eukaryotic systems available, the majority of proteins and enzymes are produced by prokaryotes. Especially if recombinant proteins shall be produced at industrial scale, bacterial hosts like *Escherichia coli* are employed. One of the reasons for the *E. coli* host system is, that the cultivation and genetics are easy to handle and appropriate techniques are established throughout most laboratories. However, in pharmaceutical protein productions the accumulation of lipopolysaccharides (LPS), generally referred as endotoxins, is a significant drawback of protein production with *E. coli*. Gram-positive bacterial hosts like *Bacillus* species do not possess LPS and thus are not limited in

their use. Further, many members of this genus have a long history of use in food fermentation processes (Hong *et al.*, 2008) and thus they are widely accepted to be applied in food industry. Their high natural protein secretion capacity is another advantage of *Bacillus* hosts systems (Freudl, 1992). The secretion of target proteins into the growth medium highly reduces costs for protein purification. In comparison to the industrially employed *Bacillus subtilis*, strains of *Bacillus megaterium* have the advantage of high stability of freely replicating plasmids and the lack of alkaline proteases (Vary, 1994). Large amounts of intact functional proteins with only little or no degradation products were obtained (Kim, 2003; von Tersch and Robbins, 1990). Thus, *B. megaterium* is a useful alternative for high yield recombinant protein production (Vary *et al.*, 2007).

I.3 *Bacillus megaterium*

The apathogenic Gram-positive *B. megaterium* (Figure I.a) was first described by de Bary more than a century ago. Its enormous size was eponymous („megat(h)erium“ -Greek for “giant beast”-) (De Bary, 1884). With up to 60 µm^3 it has a more than 100-fold larger volume than *E. coli*. Due to the size of its vegetative forms and spores *B. megaterium* is well suited for studies of cell-structure, protein localization, membrane and spore formation (Vary *et al.*, 2007). Its ability to utilize a wide variety of carbon sources enables the organism to colonize very different habitats. It was found in such diverse environments like paddy rice, dried food, seawater, sediments, fish, plants and even in honey (Vary, 1994).

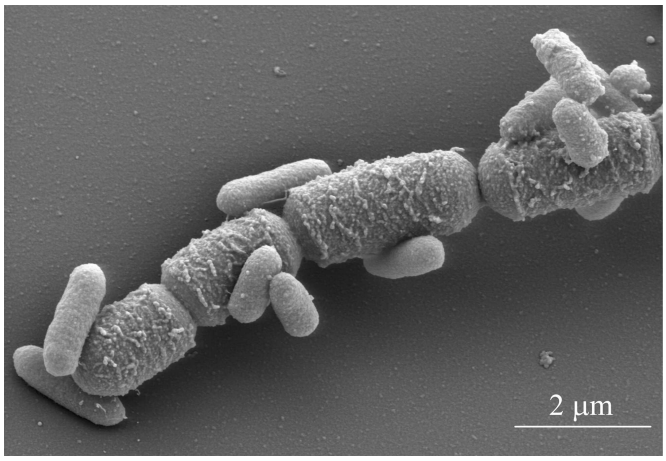

Figure I.a: Electron microscopic image of *Bacillus megaterium*.
B. megaterium (big cells) with *E. coli* (small cells) (Biedendieck *et al.*, 2007b).

Today, the genomes of 32 *Bacillus* species, including the industrial protein production hosts *B. subtilis* and *B. licheniformis* as well as several pathogenic strains of *B. anthracis* and *B. cereus* have been sequenced. Sequencing of the *B. megaterium* DSM319 genome was recently performed on behalf of the Technische Universität Braunschweig in collaboration with a team of the J. Craig Venter Institute (Rockville, MA, USA) that sequenced the genome of the *B. megaterium* QM B1551 strain in parallel. The most obvious difference between both strains are the seven plasmids, which are only present in the strain QM B1551 and comprise up to 11 % of its total cellular DNA. The size spectrum of these plasmids reaches from 5.4 kb to over 165 kb (Kieselburg *et al.*, 1984). The genome sequence of the plasmidless strain *B. megaterium* DSM319 was solved by shotgun sequencing using a combination of conventional Sanger sequencing and high-throughput pyrosequencing. The determined size of the *B. megaterium* DSM319 genome is ~5.2 Mbp and the GC-content is 38 %. Formerly, the species *B. megaterium* was placed within the *B. subtilis* group of Bacilli. Already back then, it was found to be more closely related to *B. cereus* or *B. pumilus* than to *B. subtilis* (Ash *et al.*, 1991). Latest classification of *B. megaterium* DSM319 based on 16S rRNA gene sequence analysis (Figure I.b) revealed, that this species is settled quite isolated within the group of Bacilli with its next relatives within the group of *B. cereus* (Bunk, 2010).

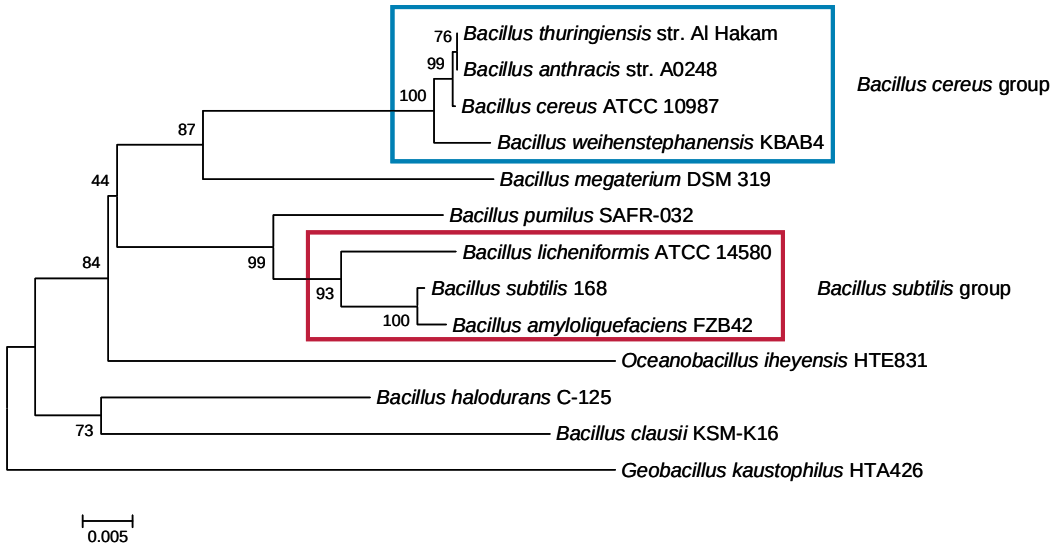

Figure I.b: Phylogenetic classification of *Bacillus megaterium* DSM319.

This classification is based on 16S rRNA gene comparison of 13 *Bacillus* species using the Neighbor-Joining method. Strains of the *B. cereus* group are marked with a blue box, species of the *B. subtilis* group with a red box (according to Bunk, 2010).

Wild-type *B. megaterium* strains were already used to generate products of industrial importance like α- and β-amylases which are used for starch modification in the baking industry (Vary, 1994). Furthermore, *B. megaterium* is utilized for the production of penicillin amidases which are essential for the synthesis of novel β-lactam antibiotics (Panbangred *et al.*, 2000). Moreover, *B. megaterium* is known for its ability to synthesize vitamin B_{12} both aerobically and anaerobically (Biedendieck *et al.*, 2010; Raux *et al.*, 1998).

I.4 The xylose-inducible gene expression system for *Bacillus megaterium*

Today, for most biotechnological applications in research and industry high yield recombinant production of proteins is desired. Hence, efficient and easy to handle genetic tools for *B. megaterium* had to be established to make it a suitable alternative host system. To facilitate recombinant approaches, Tersch and Robbins (1990) developed a transformation procedure which allowed for genetic manipulations. This enabled the introduction of recombinant plasmids into protoplasted *B. megaterium* cells (von Tersch & Robbins, 1990). These are stably maintained in *B. megaterium* over many generations even without any selective pressure. Later on, Rygus and Hillen identified and characterized the homologous xylose-inducible promoter (P_{xylA}) offering the opportunity for regulated expression of recombinant genes (Rygus & Hillen, 1991; Rygus *et al.*, 1991). In its natural context, the promoter P_{xylA} is located upstream of the operon encoding XylA (xylose isomerase), XylB (xylulokinase) and XylT (xylose permease). The gene for the repressor of P_{xylA} (*xylR*) is located upstream of the *xyl*-operon in opposite orientation. XylR is produced in active conformation and binds to a palindromic operator sequence directly downstream of P_{xylA}. The repressive effect is supposed to function by steric hampering of RNA polymerase binding, when the repressor is bound at this location. In presence of xylose, repressor binding is prevented and transcription initiation at P_{xylA} can take place. These functional elements were placed on a multi-host shuttle plasmid with the genes *xylA*, *xylB* and *xylT* replaced by a multiple cloning site (*mcs*) (Figure I.c).

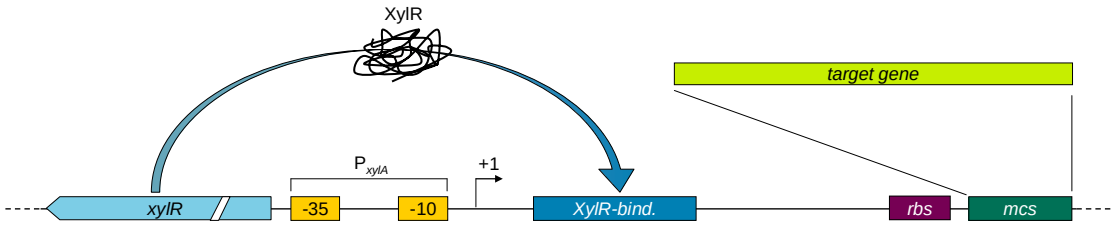

Figure I.c: Key features of the xylose-inducible gene expression system for *Bacillus megaterium*.

Elements for regulated gene expression in *B. megaterium* are the xylose-inducible promoter (P_{xylA}), the gene for the xylose repressor (XylR) and its palindromic DNA binding region (XylR-*bind.*). Binding of xylose leads to conformational changes of XylR resulting in its release from the XylR-binding region and to initiation of transcription at the +1 site. The multiple cloning site (*mcs*) for introduction of target genes is located downstream of the ribosome binding site (*rbs*).

Genetic manipulations and amplification of this plasmid can be carried out in *E. coli*, while for target gene expression the plasmid is finally transferred into *B. megaterium*. To facilitate these tasks, origins of replication (*ori*s) for both *E. coli* and *Bacillus* as well as antibiotic resistance cassettes (*amp*^{r}, *tet*^{r}) were placed on this plasmid. Later on, this system was modified and optimized regarding its applicability. An enlarged *mcs* was inserted, non functional elements of the plasmids were deleted to minimize its size and several coding sequences for protein purification tags were introduced and evaluated (Biedendieck *et al.*, 2007c; Malten *et al.*, 2005; Malten *et al.*, 2006). However, the protein yields generated by this expression system in *B. megaterium* remained limited compared to established hosts like *E. coli*.

I.5 Phage RNA polymerase-based gene expression systems

The most effective protein production systems for both prokaryotic and eukaryotic hosts are based on phage RNA polymerases (RNAPs) (Conrad *et al.*, 1996; Fuerst *et al.*, 1986; Sagawa *et al.*, 1996; Tabor & Richardson, 1985). Nearly exclusively the RNAP of the *E. coli* phage T7 is used. This bacteriophage belongs to the recently created subfamily of *Podoviridae*, that was designated *Autographivirinae* (self-transcribing) (Lavigne *et al.*, 2008). This subfamily comprises three genera of phages

which encode their own viral RNAP: the T7 *sensu strictu* group, the SP6-like and the φKMV-like phages (Lavigne *et al.*, 2008). Only few of the DNA-dependent RNAPs annotated in genomes of *Autographivirinae* were biochemically characterized. Well understood are the RNAPs of the *E. coli* phages T7 and T3 and of the *Salmonella typhimurium* phage SP6 (reviewed by Chamberlin and Ryan (1982)). All these RNAPs are single-subunit enzymes with a molecular mass of 95 to 100 kDa. They are characterized by stringent promoter recognition and a highly efficient transcription initiation. Phage RNAPs are extremely processive enzymes, which readily synthesize transcripts of thousands of nucleotides in length without dissociating from their DNA template. The T7 RNAP shows a 5-fold faster transcription than the corresponding enzyme of *E. coli* (Golomb & Chamberlin, 1974). Moreover, RNA yields of up to 1 - 2 mg mL^{-1} can be obtained in *in vitro* transcription assays (Paschal *et al.*, 2008). Consequently, these enzymes represent a powerful tool not only for protein production systems but also for the majority of *in vitro* transcription applications. When phage RNAPs are used in gene expression systems (Figure I.d), the target gene is under control of an appropriate phage RNAP promoter and terminated by a phage RNAP transcription terminator. The RNAP gene can either be located on the host chromosome, on a plasmid, or the host is infected prior to protein production with phage particles carrying the neccessary DNA encoding the RNAP. If the polymerase gene is located within the host cell, its production has to be started by addition of the appropriate inducer of the promoter applied for RNAP gene transcription. Due to the high transcription rates of phage RNAPs, even small amounts of the viral polymerase trigger the production of high amounts of target mRNA which is subsequently translated into protein by the host's translational machinery. Since phage RNAPs are not inhibited by the antibiotic rifampicin, its addition stops the host's transcriptional machinery and allows for an exclusive production of target mRNA and thus target protein (Tabor & Richardson, 1985).

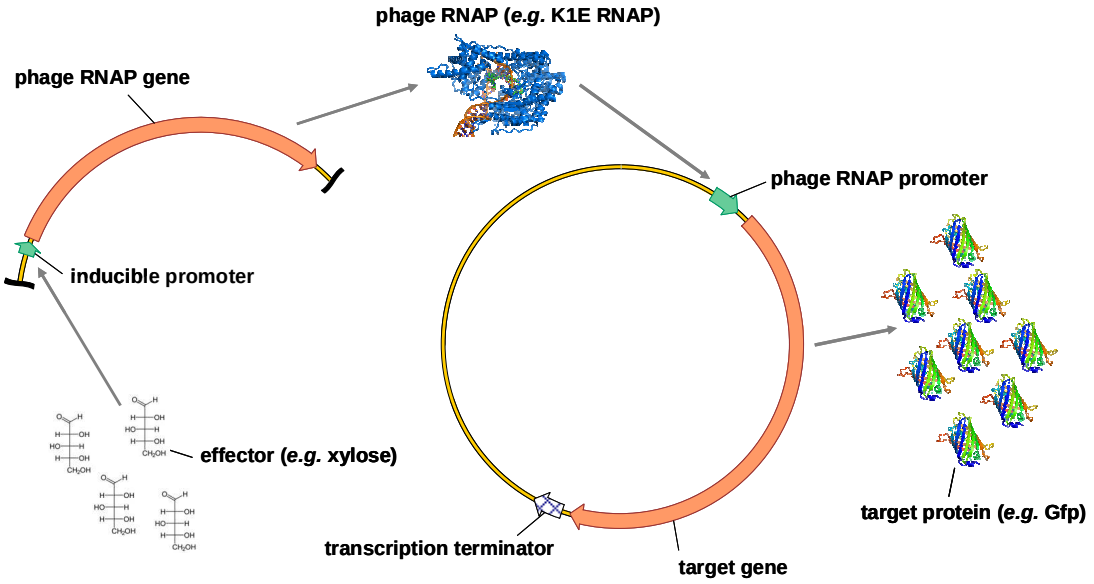

Figure I.d: Schematic image of a phage RNA polymerase-based gene expression system.
The gene encoding the phage RNAP, which is located somewhere within the host cell, is produced when the corresponding promoter is induced by its effector. Subsequently, RNAP encoding mRNA is produced and translated. The phage RNAP then immediately starts transcription of the target gene, which is under control of the cognate phage RNAP promoter. Transcription is aborted at the phage RNAP transcription terminator which is located downstream of the target gene. Rapidly, high amounts of target mRNA accumulate within the cell and are transcribed by the host's translational machinery.

I.6 Protein secretion in Bacilli

Gram-positive bacteria like *Bacillus* species have five possible locations for their proteins: 1. cytoplasm, 2. cytoplasmic membrane, 3. space between membrane and cell wall, 4. cell wall and 5. exterior. To indicate a protein's final destination, a type of "zip code" encoded as amino acid sequence is used. The importance of protein transport is underlined by the fact, that *e.g.* half of all *B. subtilis* proteins are somehow related to the cell membrane. These proteins are either inserted into, attached to, or translocated through the membrane to a destination in the cell wall or the exterior (Yamane *et al.*, 2004). The best known "zip code" is the N-terminally located secretory signal peptide which targets a protein for translocation across the cell membrane (Emanuelsson *et al.*, 2007). Five different secretion pathways are known to exist in Bacilli. These differ in specific characteristics of their leader

peptides (Tjalsma *et al.*, 2004). These are length, charge, hydrophobicity or signal peptide recognition site (Tjalsma *et al.*, 2000). Nevertheless, typically all of them contain a three domain structure (N-, H-, and C-domain). Within the N-domain at least one arginine or lysine residue is present suggested to interact with the translational machinery and the negatively charged phospholipids of the lipid bilayer. The H-domain comprises a stretch of hydrophobic amino acids often interrupted by a helix-breaking glycine or proline residue. This part adopts an α-helical conformation and integrates into the cell membrane. The signal peptidase recognition and cleavage site is located in the C-domain of the leader, which after translocation appears at the outside of the membrane. Here, the signal peptide is cleaved off by a signal peptidase and the mature protein is released. The remaining leader sequence is subsequently degraded by a signal peptide peptidase (Tjalsma *et al.*, 2004).

The major part of *B. subtilis* secretory proteins is translocated in a Sec-dependent manner (Figure I.e) (Tjalsma *et al.*, 2000). During the translation of an mRNA into the pre-polypeptide of an exoprotein components of the Sec system (Fth, HBsu and small cytoplasmic RNAs) form the signal recognition particle (SRP) and recognize the N-terminally located signal peptide. With the aid of FtsY the unfolded pre-protein is targeted to the cell membrane. The translocation machinery located in the membrane consists of at least four proteins (SecA, SecE, SecG and SecY). The transport process shows many similarities to the one of *E. coli*. SecA, the translocation motor protein transports the pre-protein through the SecYEG channel by hydrolysis of ATP. At the *trans*-side, a membrane-anchored type I signal peptidase recognizes its cleavage site, with the consensus A-S-A at the positions -3 to -1 in the C-terminus of the leader and separates the mature protein from the N-terminal signal peptide residue. The mature proteins fold autocatalytically, supported by the presence of cations or by foldases like PrsA. Cell wall associated proteases like WprA instantly degrade misfolded proteins. These enzymes represent the cells quality control system, but are paradoxically suggested to be one of the major bottlenecks in heterologous protein secretion. If finally the correctly folded proteins have reached the cell exterior, they diffuse through the cell wall and reach their destination, the environment (Tjalsma *et al.*, 2004).

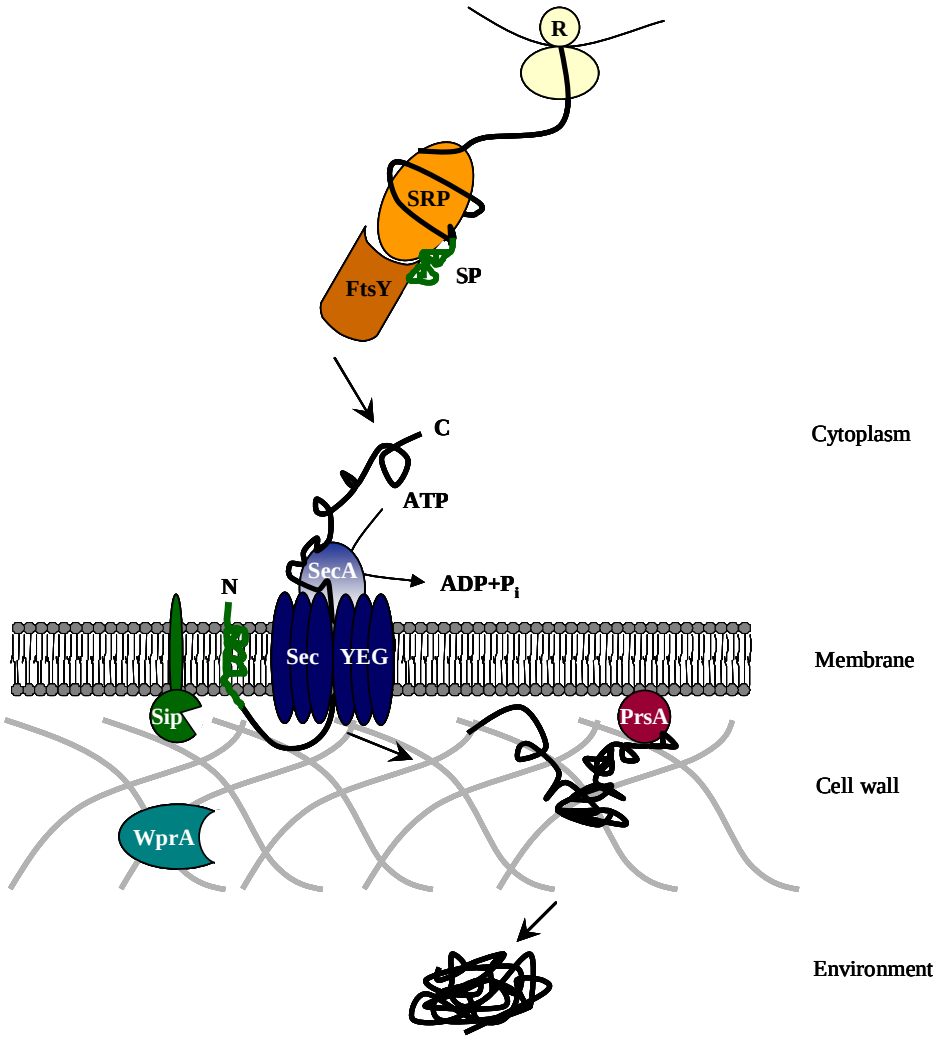

Figure I.e: Sec-dependent protein transport in *Bacillus subtilis*.
Secretory proteins (black line) are synthesized as precursor molecules with N-terminal signal peptides (SPs) at the ribosome (R). Cytosolic chaperones like the SRP-FtsY-complex bind to the pre-proteins and retain their unfolded conformation. They target the unfolded proteins to the secretion translocase in the membrane, consisting of the motor protein SecA and the three membrane proteins SecY, SecE and SecG. During or after translocation, the SPs of the pre-proteins are cleaved off by a type I signal peptidases (Sip). The autocatalytic folding of several secreted proteins is supported by the activity of extracellular chaperones like PrsA. Cell wall associated proteases like WprA are involved in quality control of exported proteins. After folding, the mature proteins passage through the cell wall and are released into the surroundings (simplified after Tjalsma *et al.* 2004).

I.7 Model proteins for protein production studies

For the evaluation and comparison of different promoters or gene expression systems model proteins are used. Their production shall appropriately display differences in promoter strength, RNA stability and translational efficiency without significant influence on the viability of the host organism. Thus, well-suited model proteins accumulate within a host cell without toxic effect and a negligible cellular stress response. Another important part is the quantification. It has to be a simple and reproducible procedure under the conditions applied for the experiment. A model protein can either be quantified by a specific colorimetric characteristic or by an enzymatic activity. In either case, it should be validated that the detected signal is not influenced by natural host proteins and / or the conditions of the experiment.

I.7.1 Green fluorescent protein (Gfp)

A variant of the green fluorescent protein (Gfp) from jellyfish *Aequoria victoria* was applied as model protein for intracellular protein production (Scholz *et al.*, 2000). The wild-type Gfp was discovered as a companion protein to the chemiluminescent protein aequorin (Shimomura *et al.*, 1962). It has a relative molecular mass of 27,000 and an excitation maximum between 395 nm and 397 nm. Wild-type Gfp fluoresces with an emission maximum of 504 nm (Tsien, 1998). Cormack *et al.* (1996) constructed a library of mutant *gfp* genes and found a double mutation resulting in a 30-times higher fluorescence intensity than the wild-type Gfp. These mutations further led to a large shift in the excitation maximum from 395 nm to 488 nm. Furthermore, its fluorescence is detectable within 8 min compared to 1 - 2 h for the wild type. Additionally, it is not harmful for the cells, independent of substrates and cofactors and provides high sensitivity. Hence, this Gfp mutant (eGfp or GFPmut1), which is commonly known to be an excellent reporter protein (Heim *et al.*, 1994; Reischer *et al.*, 2004), was used in this work. Since no other mutant protein was used in this work for simplicity it is referred to as Gfp.

I.7.2 Hydrolase of *Thermobifida fusca* (Tfh)

As model protein for production and secretion of heterologous recombinant proteins a hydrolase from *Thermobifida fusca* (Tfh) was employed. This small enzyme consists of 261 amino acid residues and has a relative molecular mass of 28,000. Tfh possesses a unique hydrolytic behavior since it acts both as an esterase and as a lipase. However, the classification as cutinase is questionable (Kleeberg *et al.*, 2005). It is able to break down aliphatic-aromatic copolyesters such as polyethylene terephtalate (PET) which are commonly regarded as being not susceptible to microbial attack (Müller *et al.*, 2005). Due to these specific features, Tfh is of considerable interest for polyester degradation (Kleeberg *et al.*, 2005). The production of Tfh in *E. coli* and first enzymatic quantification methods were established in 2006 (Dresler *et al.*). Later, Yang *et al.* (2007) designed and expressed a codon-adapted *tfh* gene in *B. megaterium* and were able to produce and secrete high amounts of this enzyme into the growth medium. An effective production and secretion of Tfh in *B. megaterium* as well as the simple and reproducible colorimetrical quantification method made this enzyme an appropriate model protein candidate.

I.8 Aim of this work

Main goal of this work was to improve the yields for both the recombinant production of heterologous proteins and their functional secretion into the growth medium. For this purpose, two different strategies were intended.

First of all, the plasmid borne gene expression system, using the native xylose-inducible promoter P_{xylA} of *B. megaterium* should systematically be optimized. Target of genetic modifications were the promoter region, the DNA sequences encoding the 5'-untranslated region (*utr*) and the ribosome binding site (*rbs*). These different DNA regions upstream of a target influence the protein production process at different stages including transcription and translation. Protein production was quantified using an enhanced green fluorescent protein variant (Gfp) as model protein. Based on this new, optimized expression system, also the utilization of the Sec-pathway for protein export should be target for further improvements. Until now, only two different signal peptides (SP_{LipA} and SP_{Pac}) were employed to facilitate the export of heterologous target proteins in *B. megaterium* (Biedendieck *et al.*, 2007a). In cooperation with Codexis Inc. (Redwood City, CA, USA) new *B. megaterium* SPs were screened and experimentally evaluated for their capability to promote protein secretion of the heterologous model hydrolase from *T. fusca* (Tfh).

Due to recurring cloning difficulties of the existing expression plasmid system in *E. coli*, a new plasmid system should be developed. Maintenance in single-copy state and "on-command" amplification prior to plasmid preparation in *E. coli* should circumvent problems related to basal expression of P_{xylA} in the cloning host. This new, so called "Ecoco" system should finally be evaluated both in *E. coli* and further for protein production as well as protein export in *B. megaterium*.

The second strategy to improve protein production rates should be based on the application of phage RNA polymerases. For this purpose, the annotated RNAP gene of the *E. coli* phage K1E should recombinantly be overproduced. After evaluation of the enzyme characteristics and determination of appropriate reaction conditions, a productive *in vitro* transcription system should be established. Finally, the new K1E RNAP was intended to be used for the construction of a new phage RNAP-based gene expression system for *B. megaterium*. Its protein production should be evaluated by production and secretion experiments using the model proteins Gfp and Tfh, respectively.

II. Materials and Methods

II.1 Instruments, chemicals and kits

II.1.1 Instruments

Instruments and technical devices used in this work are listed in Table II.a.

Table II.a: Instruments and technical devices.

Instrument	Product specification	Manufacturer
Agarose gel documentation	GelDoc	Bio-Rad
	Geldokumentation DeVision G	Decon Science Tec
Agarose gel electrophoresis	Agagel Mini	Biometra
	+ Power supply BL030623	
Autoclav	LVSA 50/70	Zirbus
	FVY/A1	Fedegari
Capillary Electrophoresis	2100 Bioanalyzer	Agilent
Cell disruption devices	French pressure cell press	Thermo
	FastPrep-24	MP Biomedicals
Centrifuges	Biofuge pico	Heraeus
	Biofuge fresco	Heraeus
	Megafuge 1.0	Heraeus
	Centrifuge 5804	Eppendorf
	MiniSpin	Eppendorf
	RC 5B Plus	Sorvall
	Avanti J-26 XP	Beckman Coulter
	Optima L-90K	Beckman Coulter
Coffee machine	Incanto	Saeco
Concentrator	Speed Vac SPD110B	Savant
	+ Refrigerated Vapour Trap RVT400	
Digital camera	Cyber shot	Sony
Distillery	Typ 2304	GFL

Instrument	Product specification	Manufacturer
DNA sequencing	ABI PRISM 310 genetic analyzer	Applied Biosystems
Fluorescence microscope	Axiovert 200M	Carl Zeiss
	+ AxioCam HR/ApoTome	
	+ Illuminator HB0/HAL100	
Gradient cycler	Tgradient	Biometra
	C1000 thermal cycler	Bio-Rad
Luminescence spectrometer	LS50B	PerkinElmer
Magnetic stirrer	MIX1	2mag
Microwave oven	MW2200	Alaska
pH determination	pH-meter CG 842	Schott
Pipettes	Reference	Eppendorf
	Research Pro	Eppendorf
Scales	SBA 52	Scaltec
	HC 52	Mettler
	BP61S	Satorius
SDS-PAGE	Mini PROTEAN II	Bio-Rad
	+ Power Pac 300	
Shaker	Bench top shaker, TR	Infors AG
	REAX 2	Heidolph
Spectrophotometer	Ultrospec 200	Biochrom
	V-550	Jasco
	+ temp. controller ETC-505T	
	NanoDrop ND-1000	NanoDrop Technologies
Thermocycler	Tpersonal	Biometra
Thermomixer	Thermomixer compact	Eppendorf
Transilluminator	Flu-O-blu	Biozym
Ultra pure water system	Synthesis A10	Millipore
Ultra sonic bath	USR57	Merck eurolab N.V.
Vortex	REAX 2000	Heidolph
	MS3 Vortexer	IKA
	Vortex-Genie 2	Scientific Industries
Water bath shaker	Aquatron	Infors AG

II.1.2 Chemicals and reagents

Chemicals and reagents used in this work are listed in Table II.b. Supplemental chemicals and reagents not listed were purchased from Difco, Fluka, GE Healthcare, Merck, Oxoid, Riedel-de-Häen, Roth and Sigma-Aldrich.

Table II.b: Chemicals and reagents.

Product	Manufacturer
Affinity material for protein purification	
Ni-NTA agarose	Qiagen
2100 Bioanalyzer kit	
RNA 6000 Nano	Agilent
RNA 6000 Pico	Agilent
Cloning kits	
CloneJet	Fermentas
TOPO-Cloning	Invitrogen
DNA purification	
QIAquick miniprep kit	Qiagen
QIAprep gel extraction kit	Qiagen
QIAquick PCR purification kit	Qiagen
RNA purification	
innuPREP RNA mini kit	Analytik Jena
DNA size standards	
GeneRuler DNA ladder mix	Fermentas
MassRuler DNA ladder mix	Fermentas
Enzymes for molecular biological applications	Biotherm
	Fermentas
	Finnzymes
	GE Healthcare
	Merck
Enzymes for molecular biological applications	New England BioLabs
	Promega Corporation
Fluorescent DNA stain	
GelStar nucleic acid gel stain	Lonza

Product		Manufacturer
Oligonucleotides		Biomers
Oligonucleotides		Metabion
Protein size standards		
	PageRuler prestained protein ladder	Fermentas
	PageRuler unstained protein ladder	Fermentas
	Protein molecular weight marker	Fermentas
Site-directed mutagenesis		
	QuikChange II	Agilent

II.2 Bacterial strains and plasmids

Bacterial strains and plasmids used in this work are listed in Table II.c and Table II.d.

Table II.c: Bacterial strains.

Strain	Description	Reference/Source
Bacillus megaterium		
DSM319	Wild type	DSMZ
MS941	Mutant of DSM319, Δ*nprM*	(Wittchen & Meinhardt, 1995)
Escherichia coli		
DH10B	F⁻ *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lac*Z ΔM15 Δ*lac*X74 *rec*A1 *end*A1 *ara*D139 Δ (*ara*, *leu*)7697 *gal*U *gal*K λ- *rps*L *nup*G /pMON14272 / pMON7124	Invitrogen
BL21(DE3)	*E. coli* B F⁻ *dcm ompT hsdS*(r_B^- m_B^-) gal λ(DE3)	Stratagene
XL10-Gold	Tet^{r}Δ(*mcrA*)*183* Δ(*mcrCB-hsdSMR-mrr*)*173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* Hte [F´ *proAB lacI*^{q}*Z*Δ*M15* Tn*10* (Tet^r) Amy Cam^r]	Stratagene

Table II.d: Plasmids.

Plasmid	Description	Reference/Source
pMM1525	Shuttle vector for xylose-inducible production and SP_{LipA} mediated secretion of target proteins; P_{xylA}-*sp*$_{lipA}$-*mcs*	(Malten *et al.*, 2006)
pHIS1525	Shuttle vector for xylose-inducible production and SP_{LipA} mediated secretion of C-terminally His_6-tagged target proteins; P_{xylA}-*sp*$_{lipA}$-*mcs*-*his*$_6$	(Malten *et al.*, 2006)
pC-HIS1622	Shuttle vector for xylose-inducible intracellular production of C-terminally His_6-tagged proteins; P_{xylA}-*mcs*-*his*$_6$-*stop*	(Biedendieck *et al.*, 2007c)
pN-HIS-TEV1622	Shuttle vector for xylose-inducible intracellular production of N-terminally His_6-tagged proteins; P_{xylA}-*his*$_6$-*tev*-*mcs*-*stop*	(Biedendieck *et al.*, 2007c)
pYYBm9	Codon optimized *tfh*-*his*$_6$ cloned into *Bgl*II and *Eag*I sites of pMM1525; P_{xylA}-*tfh*-*his*$_6$	(Yang *et al.*, 2007)
pADBm20	Shuttle vector for xylose-inducible production and SP_{Asp} mediated secretion of target proteins; P_{xylA}-*sp*$_{asp}$-*mcs*	(Biedendieck, 2006)
pRBBm26	Shuttle vector for xylose-inducible production and SP_{Pac} mediated secretion of target proteins; P_{xylA}-*sp*$_{pac}$-*strep*-*x*$_a$-*mcs*	(Biedendieck *et al.*, 2007a)
pRBBm34	Shuttle vector for xylose-inducible intracellular production of Gfp proteins; P_{xylA}-*gfp*	(Biedendieck *et al.*, 2007b)
pSTOP1622	Shuttle vector for xylose-inducible intracellular production of target proteins; P_{xylA}-*mcs*-*stop*	(Biedendieck *et al.*, 2007c)
p3STOP1622	pSTOP1622 derivative encoding three stop codons for all possible reading frames downstream of *mcs*; P_{xylA}-*mcs*-*3stop*	This work
p3STOP1623	p3STOP1622 derivative containing an additional *Pac*I restriction site; P_{xylA}-*mcs*-*3stop* (+*Pac*I)	This work
p3STOP1624	p3STOP1623 derivative containing an additional *Nhe*I restriction site; P_{xylA}-*mcs*-*3stop* (+*Pac*I/*Nhe*I)	This work
p3STOP1622-*gfp*	*gfp* cloned into *Bgl*II and *Eag*I sites of p3STOP1622; P_{xylA}-*gfp*	This work
p3STOP1623-*gfp*	*gfp* cloned into *Bgl*II and *Eag*I sites of p3STOP1623; P_{xylA}-*gfp* (+*Pac*I)	This work

Plasmid	Description	Reference/Source
p3STOP1624-*gfp*	*gfp* cloned into *Bgl*II and *Eag*I sites of p3STOP1624; P_{xylA}-*gfp* (+*Pac*I/*Nhe*I)	This work
pKMBm1	Annealed oligos -10^{+}-for and -10^{+}-rev inserted into *Pac*I and *Nhe*I site of p3STOP1624; P_{xylA}-(-10^{+})-*mcs*	This work
pSSBm40	Annealed oligos *utr*$^{+}$-for and *utr*$^{+}$-rev inserted into *Pac*I and *Nhe*I site of p3STOP1624; P_{xylA}-(*utr*$^{+}$)-*mcs*	This work
pSSBm44	Annealed oligos *rbs*$^{+}$-for and *rbs*$^{+}$-rev inserted into *Nhe*I and *Bsr*GI site of p3STOP1624; P_{xylA}-(*rbs*$^{+}$)-*mcs*	This work
pKMBm9	*gfp* cloned into *Bgl*II and *Eag*I sites of pKMBm1; P_{xylA}-(-10^{+})-*gfp*	This work
pSSBm46	*gfp* cloned into *Bgl*II and *Eag*I sites of pSSBm40; P_{xylA}-(*utr*$^{+}$)-*gfp*	This work
pSSBm50	*gfp* cloned into *Bgl*II and *Eag*I sites of pSSBm44; P_{xylA}-(*rbs*$^{+}$)-*gfp*	This work
pSSBm84	p3STOP1624-*gfp* derivative encoding an optimized -35 region of P_{xylA}; P_{xylA}-(-35^{+})-*gfp*	This work
pSSBm78	pSSBm50 derivative encoding an optimized -35 region of P_{xylA}; P_{xylA}-(-35^{+}+*rbs*$^{+}$)-*gfp*	This work
pSSBm74	Annealed oligos -10^{+}-for and -10^{+}-rev inserted into *Pac*I and *Nhe*I sites of pSSBm44; P_{xylA}-(-10^{+}+*rbs*$^{+}$)-*mcs*	This work
pKMBm4	Annealed oligos *utr*$^{+}$-for and *utr*$^{+}$-rev inserted into *Pac*I and *Nhe*I sites of pSSBm44; P_{xylA}-(*utr*$^{+}$+ *rbs*$^{+}$)-*mcs*	This work
pSSBm76	*gfp* cloned into *Bgl*II and *Eag*I sites of pSSBm74; P_{xylA}-(-10^{+}+*rbs*$^{+}$)-*gfp*	This work
pKMBm10	*gfp* cloned into *Bgl*II and *Eag*I sites of pKMBm4; P_{xylA}-(*utr*$^{+}$+*rbs*$^{+}$)-*gfp*	This work
pSSBm81	Optimized -35 region of P_{xylA} cloned into pSSBm76; P_{xylA}-(-10^{+}+ -35^{+}+*rbs*$^{+}$)-*gfp*	This work
pSSBm85	pSSBm78 derivative lacking *Nhe*I site; P_{xylA}-(-35^{+}+*rbs*$^{+}$)-*gfp* (-*Nhe*I)	This work
pC-HIS1623hp	*mcs-his*$_6$ of pC-HIS1622 inserted into pSSBm85; P_{xylA}-(-35^{+}+*rbs*$^{+}$)-*mcs-his*$_6$-*stop*	This work
pN-HIS-TEV1623hp	*mcs-his*$_6$-*tev* of pN-HIS-TEV1622 inserted into pSSBm85; P_{xylA}-(-35^{+}+*rbs*$^{+}$)-*his*$_6$-*tev-mcs-stop*	This work

Plasmid	Description	Reference/Source
p3STOP1623hp	*mcs* of p3STOP1624 inserted into pSSBm85; P_{xylA}-(-35$^+$+*rbs*$^+$)-*mcs-3stop*	This work
pSSBm22	sp_{yngK} cloned into *Bsr*GI and *Bgl*II sites of pMM1525; P_{xylA}-sp_{yngK}-*mcs*	This work
pSSBm23	sp_{vpr} cloned into *Bsr*GI and *Bgl*II sites of pMM1525; P_{xylA}-sp_{vpr}-*mcs*	This work
pSSBm24	sp_{nprM} cloned into *Bsr*GI and *Bgl*II sites of pMM1525; P_{xylA}-sp_{nprM}-*mcs*	This work
pSSBm25	sp_{yocH} cloned into *Bsr*GI and *Bgl*II sites of pMM1525; P_{xylA}-sp_{yocH}-*mcs*	This work
pSSBm27	Annealed oligos sp^+_{yngK}-for and sp^+_{yngK}-rev cloned into *Bsr*GI and *Bgl*II sites of pMM1525; P_{xylA}-sp^+_{yngK}-*mcs*	This work
pSSBm28	*tfh-his*$_6$ cloned into *Eag*I and *Spe*I sites of pSSBm22; P_{xylA}-sp_{yngK}-*tfh-his*$_6$	This work
pSSBm29	*tfh-his*$_6$ cloned into *Eag*I and *Spe*I sites of pSSBm23; P_{xylA}-sp_{vpr}-*tfh-his*$_6$	This work
pSSBm30	*tfh-his*$_6$ cloned into *Eag*I and *Spe*I sites of pSSBm24; P_{xylA}-sp_{nprM}-*tfh-his*$_6$	This work
pSSBm31	*tfh-his*$_6$ cloned into *Eag*I and *Spe*I sites of pSSBm25; P_{xylA}-sp_{yocH}-*tfh-his*$_6$	This work
pSSBm33	*tfh-his*$_6$ cloned into *Eag*I and *Spe*I sites of pSSBm27; P_{xylA}-sp^+_{yngK}-*tfh-his*$_6$	This work
pSSBm34	sp_{asp} cloned into *Spe*I and *Xho*I sites of pSSBm28; P_{xylA}-sp_{asp}-*tfh-his*$_6$	This work
pSSBm35	sp_{pac} cloned into *Spe*I and *Xho*I sites of pSSBm28; P_{xylA}-sp_{pac}-*tfh-his*$_6$	This work
pSSBm94	sp_{yngK}-*tfh-his*$_6$ subcloned from pSSBm28 into *Bsr*GI and *Pst*I sites of p3STOP1623hp; P_{xylA}-(-35$^+$+*rbs*$^+$)-sp_{yngK}-*tfh-his*$_6$	This work
pSSBm95	sp_{vpr}-*tfh-his*$_6$ subcloned from pSSBm29 into *Bsr*GI and *Pst*I sites of p3STOP1623hp; P_{xylA}-(-35$^+$+*rbs*$^+$)-sp_{vpr}-*tfh-his*$_6$	This work
pSSBm96	sp_{nprM}-*tfh-his*$_6$ subcloned from pSSBm30 into *Bsr*GI and *Pst*I sites of p3STOP1623hp; P_{xylA}-(-35$^+$+*rbs*$^+$)-sp_{nprM}-*tfh-his*$_6$	This work
pSSBm97	sp_{yocH}-*tfh-his*$_6$ subcloned from pSSBm31 into *Bsr*GI and *Pst*I sites of p3STOP1623hp; P_{xylA}-(-35$^+$+*rbs*$^+$)-sp_{yocH}-*tfh-his*$_6$	This work

Plasmid	Description	Reference/Source
pSSBm98	sp^{+}_{yngK}-*tfh*-his_6 subcloned from pSSBm33 into *Bsr*GI and *Pst*I sites of p3STOP1623hp; P_{xylA}-(-35^{+}+rbs^{+})-sp^{+}_{yngK}-*tfh*-his_6	This work
pSSBm99	sp_{asp}-*tfh*-his_6 subcloned from pSSBm34 into *Bsr*GI and *Pst*I sites of p3STOP1623hp; P_{xylA}-(-35^{+}+rbs^{+})-sp_{asp}-*tfh*-his_6	This work
pSSBm100	sp_{pac}-*tfh*-his_6 subcloned from pSSBm35 into *Bsr*GI and *Pst*I sites of p3STOP1623hp; P_{xylA}-(-35^{+}+rbs^{+})-sp_{pac}-*tfh*-his_6	This work
pSSBm101	sp_{lipA}-*tfh*-his_6 subcloned from pYYBm9 into *Bsr*GI and *Pst*I sites of p3STOP1623hp; P_{xylA}-(-35^{+}+rbs^{+})-sp_{lipA}-*tfh*-his_6	This work
pSP_{YngK}-hp	*mcs*-his_6 of pC-HIS1623hp inserted into *Spe*I and *Pst*I sites of pSSBm94; P_{xylA}-(-35^{+}+rbs^{+})-sp_{yngK}-*mcs*-his_6	This work
pSP_{Vpr}-hp	*mcs*-his_6 of pC-HIS1623hp inserted into *Spe*I and *Pst*I sites of pSSBm95; P_{xylA}-(-35^{+}+rbs^{+})-sp_{vpr}-*mcs*-his_6	This work
pSP_{NprM}-hp	*mcs*-his_6 of pC-HIS1623hp inserted into *Spe*I and *Pst*I sites of pSSBm96; P_{xylA}-(-35^{+}+rbs^{+})-sp_{nprM}-*mcs*-his_6	This work
pSP_{YocH}-hp	*mcs*-his_6 of pC-HIS1623hp inserted into *Spe*I and *Pst*I sites of pSSBm97; P_{xylA}-(-35^{+}+rbs^{+})-sp_{yocH}-*mcs*-his_6	This work
pSP^{+}_{YngK}-hp	*mcs*-his_6 of pC-HIS1623hp inserted into *Spe*I and *Pst*I sites of pSSBm98; P_{xylA}-(-35^{+}+rbs^{+})-sp^{+}_{yngK}-*mcs*-his_6	This work
pSP_{Asp}-hp	*mcs*-his_6 of pC-HIS1623hp inserted into *Spe*I and *Pst*I sites of pSSBm99; P_{xylA}-(-35^{+}+rbs^{+})-sp_{asp}-*mcs*-his_6	This work
pSP_{Pac}-hp	*mcs*-his_6 of pC-HIS1623hp inserted into *Spe*I and *Pst*I sites of pSSBm100; P_{xylA}-(-35^{+}+rbs^{+})-sp_{pac}-*mcs*-his_6	This work
pSP_{LipA}-hp	sp_{lipA}-*mcs*-his_6 of pC-HIS1525 inserted into *Bsr*GI and *Pst*I sites of pSSBm101; P_{xylA}-(-35^{+}+rbs^{+})-sp_{lipA}-*mcs*-his_6	This work
pIH41	pBAC / *oriV* derivate used as PCR template for amplification of elements for "on command" plasmid copy number control	(Heinemann, 2007)
pBAD33	*E. coli* plasmid for tight regulation and high-level expression driven by the arabinose inducible pBAD promoter	(Guzman *et al.*, 1995)

Plasmid	Description	Reference/Source
pEc-*trfA*	*trfA* of pIH41 inserted into *Bam*HI and *Sph*I sites of pBAD33; pBAD-*trfA*	This work
pBmEcoco	genetic elements for "on command" amplification of plasmid copy number of pBAC / *oriV* were inserted into p3STOP1623hp via *Aat*II and *Sph*I; the *Spe*I site (in *parB*) and the *Xho*I site (downstream of *oriV*) were deleted; P_{xylA}-(-35^{+}+*rbs*$^{+}$)-*mcs*	This work
pBmEcoco-*gfp*	*gfp* of pSSBm85 inserted into pBmEcoco via *Spe*I and *Sph*I; P_{xylA}-(-35^{+}+*rbs*$^{+}$)-*gfp*	This work
pBmEcoco-*tfh*	sp_{yocH}-*tfh* of pSSBm97 inserted into pBmEcoco via *Pac*I and *Sph*I; P_{xylA}-(-35^{+}+*rbs*$^{+}$)-sp_{yocH}-*tfh*	This work
pET-32a(+)	*E. coli* plasmid for T7-RNAP driven production of His_6-tagged target proteins	Merck KGaA
pET-32a(+)-K1EP	*k1ep* cloned into pET-32a(+) via *Eco*RI and *Eag*I	This work
pUCP20t	mobilizable derivative of broad host range vector pUCP20	(Schweizer *et al.*, 1996)
pUC18$^{(*)}$	*E. coli* high-copy number plasmid encoding a N-terminal part of *lacZ* under control of the IPTG inducible P_{lac}-promoter capable to α-complement *lacZ*ΔM15 mutations $^{(*)}$ Plasmid contains the frame shift mutation in *lacZ* described by Lobet *et al.* (1989)	Agilent
pSSEc4	pUC18$^{(*)}$ derivative additionally coding for the K1E-1 promoter	This work
pK1Eivt	pUC18-based plasmid, enabling blue/white selection and encoding a K1E-RNAP promoter for *in vitro* transcription; P_{K1EP}-*mcs*	This work
pNIAysic1	pSP64 Poly(A) vector derivative carrying the 630 bp yeast derived control transcript YPL075W_16_412249_415357_INTRON_9_759	(Carter *et al.*, 2005)
pK1Eivt-NIAysic1	NIAysic1 insert cloned into *Eco*RI and *Hind*III sites of pK1Eivt	This work
pMGBm19	Shuttle vector using the *ori* of pBM100 for replication and P_{xylA} for xylose-inducible protein production in *B. megaterium*	(Gamer *et al.*, 2009)
pK1E-RNAP	*k1ep* cloned into pMGBm19 via *Spe*I and *Sac*I; P_{xylA}-*k1ep*	This work

Plasmid	Description	Reference/Source
pSP6-RNAP	*sp6p* cloned into pMGBm19 via *Sac*I and *BamH*I; P_{xylA}-*sp6p*	This work
pP_{K1E-1}	oligos P_{K1E-1}-for and P_{K1E-1}-rev inserted via *Afl*II and *Bsr*GI into p3STOP1624; P_{K1E-1}-*mcs*	This work
pP_{K1E-2}	oligos P_{K1E-2}-for and P_{K1E-2}-rev inserted via *Afl*II and *Bsr*GI into p3STOP1624; P_{K1E-2}-*mcs*	This work
pP_{SP6}	oligos P_{SP6}-for and P_{SP6}-rev inserted via *Afl*II and *BsrG*I into p3STOP1624; P_{SP6}-*mcs*	This work
$pP_{K1E-1+t}$	oligos term-for and term-rev inserted via *Sph*I and *Age*I into pP_{K1E-1}; P_{K1E-1}-*mcs*-term	This work
$pP_{K1E-2+t}$	oligos term-for and term-rev inserted via *Sph*I and *Age*I into pP_{K1E-2}; P_{K1E-2}-*mcs*-term	This work
pP_{SP6+t}	oligos term-for and term-rev inserted via *Sph*I and *Age*I into $pP_{SP6;}$ P_{SP6}-*mcs*-term	This work
pP_{SP6+t}-*gfp*	*gfp* cloned into *Spe*I and *Sph*I sites of pP_{SP6+t}; P_{SP6}-*gfp*-t	This work
pP_{K1E-1}-*gfp*	*gfp* cloned into *Spe*I and *Sph*I sites of pP_{K1E-1}; P_{K1E-1}-*gfp*	This work
$pP_{K1E-1+t}$-*gfp*	*gfp* cloned into *Spe*I and *Sph*I sites of $pP_{K1E-1+t}$; P_{K1E-1}-*gfp*-t	This work
$pP_{K1E-2+t}$-*gfp*	*gfp* cloned into *Spe*I and *Sph*I sites of $pP_{K1E-2+t}$; P_{K1E-2}-*gfp*-t	This work
pT7-RNAP	*B. megaterium* plasmid harboring the gene for T7 RNAP under control of the xylose-inducible promoter	(Gamer *et al.*, 2009)
pP_{T7}	*B. megaterium* plasmid for insertion of target genes under control of the T7 promoter	(Gamer *et al.*, 2009)
pP_{T7}-*gfp*	*B. megaterium* plasmid encoding Gfp under control of the T7 promoter	(Gamer *et al.*, 2009)
$pP_{K1E-1+t}$-*tfh*	sp_{yocH}-*tfh*-his_6 of pSSBm97 cloned into *Bsr*GI and *Sph*I sites of $pP_{K1E-1+t}$; P_{K1E-1}-*tfh*-his_6-t	This work
pP_{SP6+t}-*tfh*	sp_{yocH}-*tfh*-his_6 of pSSBm97 cloned into *Bsr*GI and *Sph*I sites of pP_{SP6+t}; P_{SP6}-*tfh*-his_6-t	This work
pP_{T7}-*tfh*	sp_{yocH}-*tfh*-his_6 of pSSBm97 cloned into *Bsr*GI and *Sph*I sites of pP_{T7}; P_{T7}-*tfh*-his_6-t	This work

II.2.1 Plasmid construction

II.2.1.1 Vectors used for the optimization of the P_{xylA} system

The basic expression plasmid of this investigation p3STOP1622, containing one stop-codon in each possible reading frame downstream of the *mcs*, was constructed by site-directed mutagenesis of pSTOP1622 (Biedendieck *et al.*, 2007c). The oligonucleotides QC-p3STOP1622-for and QC-p3STOP1622-rev were used as primers for the PCR reaction to introduce two additional stop-codons downstream of the *mcs* into pSTOP1622. Starting from the resulting p3STOP1622, two unique restriction sites were introduced to allow for simple genetic modification of the DNA region between P_{xylA} and the *mcs*. For this purpose, a *Pac*I-site was inserted between the -10 and -35 region of P_{xylA} by site-directed mutagenesis using the primer pair QC-p3STOP1623-for and QC-p3STOP1623-rev resulting in p3STOP1623. Afterwards, a *Nhe*I site was introduced by replacing the DNA fragment between the *Pac*I and *Bsr*GI restriction sites with the synthetic oligonucleotide pair p3STOP1624-for and p3STOP1624-rev resulting in p3STOP1624. The *gfp* gene was amplified by PCR from pRBBm34 using the primers *gfp*-for and *gfp*-rev (Biedendieck *et al.*, 2007c). The gene, here referred to as *gfp* differed from the *egfp* gene employed by Biedendieck *et al.* (2007c) by a single-base exchange, which leads to an amino acid exchange at the C-terminus of the protein without obvious influence on the protein function. Insertion of *gfp* into the *mcs* of p3STOP1622, p3STOP1623 and p3STOP1624 via the corresponding *Bgl*II and *Eag*I restriction sites led to p3STOP1622-*gfp*, p3STOP1623-*gfp* and p3STOP1624-*gfp*, respectively.

The oligonucleotide pairs -10^+-for - -10^+-rev, *utr*$^+$-for - *utr*$^+$-rev and *rbs*$^+$-for - *rbs*$^+$-rev, carrying optimized genetic elements for transcription and translation were introduced into p3STOP1624 after *Pac*I - *Nhe*I or *Nhe*I - *Bsr*GI digestion. The resulting plasmids were named pKMBm1 (optimized -10 region), pSSBm40 (optimized *utr*) and pSSBm44 (optimized *rbs*). The *gfp* gene was introduced into these plasmids as described above resulting in pKMBm9, pSSBm46 and pSSBm50. The plasmids pSSBm84 and pSSBm78 with an optimized -35 region for the P_{xylA} promoter were generated by site-directed mutagenesis. For the necessary PCR reactions the primers QC-35^+-for and QC-35^+-rev were used in combination with p3STOP1624-*gfp* and pSSBm50 as template. Fusion of the *Xho*I - *Nhe*I fragment from pKMBm1

harboring an optimized -10 region for the P_{xylA} promoter with the vector pSSBm44 containing an optimized ribosome-binding site (*rbs*$^+$) resulted in plasmid pSSBm74. Ligation of the oligos *utr*$^+$-for and *utr*$^+$-rev with the *Nhe*I and *Bsr*GI cut pSSBm44 vector resulted in pKMBm4. The *gfp* gene was inserted into the *mcs* of both plasmids pSSBm74 and pKMBm4 as described above resulting in pSSBm76 and pKMBm10. Fusion of the *Aat*II - *Pac*I fragments derived from pSSBm78 and from pSSBm76 resulted in the plasmid pSSBm81 which contains both optimized promoter elements (-10$^+$ and -35$^+$) in combination with the optimized *rbs*$^+$ and the *gfp* gene. Site-directed mutagenesis of pSSBm78 using the primer pair QC-Δ*Nhe*I-for and QC-Δ*Nhe*I-rev led to the *Nhe*I lacking plasmid pSSBm85.

The *mcs* and affinity tag encoding regions were subcloned from pC-HIS1622, pN-HIS-TEV1622 (Biedendieck *et al.*, 2007c) and p3STOP1624 into pSSBm85 using the restriction sites *Bsr*GI and *Pst*I. Thereby, the high performance (hp) plasmids pC-HIS1623hp, pN-HIS-TEV1623hp and p3STOP1623hp were generated.

Four signal peptide (SP) coding regions for SP_{YngK}, SP_{Vpr}, SP_{NprM} and SP_{YocH} were amplified by PCR from genomic *B. megaterium* DSM319 DNA using the primer pairs *sp*$_{yngK}$-for - *sp*$_{yngK}$-rev, *sp*$_{vpr}$-for - *sp*$_{vpr}$-rev, *sp*$_{nprM}$-for - *sp*$_{nprM}$-rev and *sp*$_{yocH}$-for - *sp*$_{yocH}$-rev, respectively. The resulting PCR fragments and the annealed oligonucleotides *sp*$^+_{yngK}$-for and *sp*$^+_{yngK}$-rev, which encode a codon adapted SP_{YngK}, were inserted into pMM1525 via the *Eag*I and *Spe*I sites, resulting in the plasmids pSSBm22, pSSBm23, pSSBm24, pSSBm25 and pSSBm27. Using vector pYYBm9 as template, the codon-optimized gene of a *T. fusca* hydrolase (Tfh) fused to a His_6-tag was amplified via PCR using the primers *tfh*-for and *tfh*-rev. Plasmids pSSBm22, pSSBm23, pSSBm24, pSSBm25 and pSSBm27 were cut with *Eag*I and *Spe*I and subsequently ligated with the amplified *tfh-his*$_6$. The resulting vectors were named pSSBm28, pSSBm29, pSSBm30, pSSBm31 and pSSBm33. DNA fragments containing coding sequences for either an artificial signal peptide (SP_{Asp}) or the signal peptide from *B. megaterium* penicillin amidase SP_{Pac} were amplified by PCR using the templates pADBm20 and pRBBm26 and the primer pairs *sp*$_{asp/pac}$-for - *sp*$_{asp}$-rev and *sp*$_{asp/pac}$-for - *sp*$_{pac}$-rev, respectively. Both PCR products were inserted into pSSBm28 via the restriction sites *Spe*I and *Xho*I creating the plasmids pSSBm34 and pSSBm35. The DNA regions encoding the different SP-Tfh-His_6 fusions were subcloned from the various basic expression plasmids into the high performance plasmid p3STOP1623hp utilizing the restriction sites *Bsr*GI and *Pst*I. Thereby, the

plasmid series pSSBm94 to pSSBm101 was generated.

Insertion of *mcs* and His_6-Tag coding DNA region derived from pC-HIS1623hp into the vectors pSSBm94 to pSSBm100 via the *Spe*I - *Pst*I restriction sites finally resulted in new high performance (hp) plasmids for extracellular protein production. According to the encoded signal peptides these plasmids were named pSP_{YngK}-hp, pSP_{Vpr}-hp, pSP_{NprM}-hp, pSP_{YocH}-hp, pSP^{+}_{YngK}-hp, pSP_{Asp}-hp and pSP_{Pac}-hp. Plasmid pSP_{LipA}-hp was constructed by insertion of sp_{lipA}-*mcs*-his_6 of pHIS1525 (Malten *et al.*, 2006) into the high performance plasmid derivate pSSBm101.

II.2.1.2 Plasmids constructed for the Ecoco system

The gene encoding the up-mutant of TrfA was amplified via PCR from the pBAC / *oriV* plamid derivate pIH41 using the primers *trfA*-for and *trfA*-rev. This PCR fragment was inserted into pBAD33 via *Bam*HI and *Sph*I resulting in pEc-*trfA*. An additional base was found in the sequence encoding the later 5' untranslated mRNA region of the *trfA* transcript. An obvious effect of this mutation was not detected. The elements for "on command" regulation of plasmid copy number in *E. coli* (with exception of the *trfA* gene) where amplified via PCR using the template pIH41 and the primer pair coco-for - coco-rev. This fragment was inserted into the high-performance *B. megaterium* plasmid p3STOP1623hp via *Sph*I and *Aat*II. In two subsequent site-directed mutagenesis of this plasmid, excess restriction sites were deleted. For this purpose, the primer pairs QC-*Spe*I-for - QC-*Spe*I-rev and QC-*Xho*I-for - QC-*Xho*I-rev were employed. Thus, finally the plasmid pBmEcoco was created. Excision of the *gfp* gene of pSSBm85 by *Spe*I and *Sph*I and insertion into pBmEcoco led to pBmEcoco-*gfp*. The coding region of the secretory leader peptide SP_{yocH} and the hydrolase gene *tfh* of *T. fusca* were excised from pSSBm97 by *Pac*I and *Sph*I. Insertion of this fragment into pBmEcoco resulted in pBmEcoco-*tfh*.

II.2.1.3 Plasmids developed for the phage RNA polymerase systems

Gene "*7*" of enterobacteriophage K1E (in this work named *k1ep*) was amplified from genomic DNA (NCBI GeneBank Accession No: AM084415.1) (gift of Dr. Martina Mühlenhoff and Dr. Rita Gerardy-Schahn, MH Hannover, Germany) by PCR using

the primers K1EP_a-for and K1EP_a-rev. The PCR product encoding the K1E phage RNAP was inserted into the expression plasmid pET-32a(+) via the *Eco*RI and *Eag*I restriction sites resulting in pET-32a(+)-K1EP.

A PCR using the primers pSSEc4-for - pSSEc4-rev and pUCP20t as template led to a plasmid fragment encoding the K1E RNAP promoter K1E-1 (atttactggacactatagaa). Insertion of this PCR product into pUC18$^{(*)}$ via *Eco*RI and *Sap*I resulted in pSSEc4. pUC18$^{(*)}$ is the pUC18 derivative described by Lobet *et al.* (1989). It carries a frame-shift mutation within the second codon of the *lacZ* gene fragment. This reduces production of LacZ significantly and thus hampers blue / white screening (Lobet *et al.*, 1989). Insertion of the hybridized double stranded oligonucleotides pK1Eivt-for and pK1Eivt-rev into pSSEc4 via the *Eco*RI and *Xma*I restriction sites resulted in the basic *in vitro* transcription vector pK1Eivt. By using this cloning approach the frame-shift mutation present in pUC18$^{(*)}$ was eliminated and a functional *lacZ* fragment was reconstituted. The DNA sequence encoding an exogenous RNA control derived from the yeast genome (YPL075W_16_412249_415357_INTRON_9_759) was amplified by PCR employing the primers NIAysic1-for and NIAysic1-rev and the template pNIAysic1 (Carter *et al.*, 2005). Introducing this target for *in vitro* transcription into pK1Eivt via *Eco*RI and *Hind*III led to pK1Eivt-NIAysic1.

The gene *k1ep* and gene "*8*" of the SP6 phage genome (in this work named *sp6p*) were amplified from genomic K1E and SP6 DNA (NCBI GeneBank Accession No: AY288927) (gift of Dr. Graham Hatfull and Dr. Roger Hendrix, University of Pittsburgh, PA, USA) using the primer pairs K1EP_b-for - K1EP_b-rev and SP6P-for - SP6P-rev, respectively. Amplified genes were placed into pMGBm19 under control of P_{xylA} via *Spe*I and *Sac*I (*k1ep*) or *Sac*I and *Bam*HI restriction sites (*sp6p*). The resulting vectors were named pK1E-RNAP and pSP6-RNAP, respectively.

Oligonucleotide pairs P_{K1E-1}-for - P_{K1E-1}-rev, P_{K1E-2}-for - P_{K1E-2}-rev and P_{SP6}-for - P_{SP6}-rev coding for the indicated RNAP promoter and the DNA fragments upstream of the associated genes "*31*" or "*5*" of bacteriophage K1E or gene "*49*" of SP6 phage were synthesized. The ribosome binding sites were changed into optimal ones for *B. megaterium* (Stammen *et al.*, 2010a). The hybridized and double stranded oligonucleotides were introduced into p3STOP1624 individually via the restriction sites *Afl*II and *Bsr*GI creating the plasmids pP_{K1E-1}, pP_{K1E-2} and pP_{SP6}, respectively. Putative RNAP transcription terminators of the K1E phage (NCBI GeneBank Accession No AM084415.1: 25,713 - 25,744) encoded by the oligonucleotide pair

term-for and term-rev were inserted via *Sph*I and *Age*I restriction sites into pP_{K1E-1}, pP_{K1E-2} and pP_{SP6}. Consequently, the developed vectors were named $pP_{K1E-1+t}$, $pP_{K1E-2+t}$ and pP_{SP6+t}. The *gfp* gene was amplified by PCR using pSSBm85 as template and the primer *gfp*-for and *gfp*-rev. The *Spe*I and *Sph*I cut PCR product was introduced into pP_{SP6+t}, pP_{K1E-1}, $pP_{K1E-1+t}$ and $pP_{K1E-2+t}$ generating the plasmids pP_{SP6+t}-*gfp*, pP_{K1E-1}-*gfp*, $pP_{K1E-1+t}$-*gfp* and $pP_{K1E-2+t}$-*gfp*. A sp_{yocH}-*tfh* containing fragment was cut out from pSSBm97 employing *Bsr*GI and *Sph*I. Subsequently, the inserted was ligated into appropriate cut pP_{SP6+t}, pP_{K1E-1} and pP_{T7} plasmids resulting in pP_{SP6+t}-*tfh*, $pP_{K1E-1+t}$-*tfh* and pP_{T7}-*tfh*, respectively.

II.3 Growth media and media additives

II.3.1 Growth media

As standard medium for growth of all bacterial strains, Luria Bertani (LB) medium (Sambrook *et al.*, 2001) was used unless indicated otherwise. For solid media, 1.5 % (w/v) agar-agar was added prior to sterilization.

Luria-Bertani Broth (LB medium)

Tryptone	1.0	% (w/v)
Yeast extract	0.5	% (w/v)
NaCl	0.5	% (w/v)

Recombinant protein production experiments were performed in LB medium or Fermentation medium.

Fermentation medium

$MgSO_4$ x 7 H_2O	300.0	mg L^{-1}
$(NH_4)_2SO_4$	25.0	g L^{-1}
Prepared as 5-fold stock solution and sterilized.		
KH_2PO_4	3.52	g L^{-1}
Na_2HPO_4 x 2 H_2O	6.62	g L^{-1}
Prepared as 10-fold stock solution and sterilized.		
$MnCl_2$ x 4H_2O	80.0	mg L^{-1}

Fermentation medium

$CaCl_2$ x 2 H_2O	106.0	mg L^{-1}
$FeSO_4$ x 7 H_2O	5.0	mg L^{-1}
$(NH_4)_6Mo_7O_{24}$ x 4 H_2O	4.0	mg L^{-1}
$CoCl_2$	2.2	mg L^{-1}

Prepared as 500-fold stock solution and sterilized by filtration (pore size 0.2 µm).

Fructose	5.0	g L^{-1}

Prepared as 50-fold stock solution and sterilized by filtration (pore size 0.2 µm).

Prior to use, appropriate amounts of the stock solutions were mixed and filled up with autoclaved dH_20 to the desired volume.

II.3.2 Media additives

Antibiotics and other additives were prepared as concentrated stock solutions, sterilized by filtration (pore size 0.2 µm) and added into the medium after autoclaving. Solvents and concentrations are summarized in Table II.e.

Table II.e: Media additives.

Additive	Bacterial strain	Stock solution	Final concentration
Carbenicillin	*E. coli*	100.0 mg mL^{-1} in dH_2O	50.0 µg mL^{-1} 100.0 µg mL^{-1}
Chloramphenicol	*E. coli*	34.0 mg mL^{-1} in 70 % (v/v) ethanol	17.0 µg mL^{-1} 34.0 µg mL^{-1}
	B. megaterium	4.5 mg mL^{-1} in 70 % (v/v) ethanol	4.5 µg mL^{-1}
Tetracyclin	*B. megaterium*	5.0 mg mL^{-1} in 70 % (v/v) ethanol	10.0 µg mL^{-1}
Xylose	*B. megaterium*	50.0 % (w/v) in dH_2O	0.5 % (w/v)
Arabinose	*E. coli*	2.0 % (w/v) in dH_2O	0.02 % (w/v)
IPTG	*E. coli*	100.0 mM in dH_2O	0.1 mM

II.4 Microbiological techniques

II.4.1 Sterilization

All media were vapor-sterilized at 121 °C and 1 bar overpressure for 20 min. Other substances and solutions were either vapor-sterilized or, if temperature sensitive, sterilized by filtration (pore width 0.2 µm).

II.4.2 Plate cultures

Bacteria were taken from glycerol stocks (section II.4.6) with an inoculation loop and plated onto LB medium agar plates. When required, antibiotics or other additives were added to the medium. Agar plates were incubated overnight at 30 °C or 37 °C.

II.4.3 Liquid cultures of *Escherichia coli*

Liquid cultures were inoculated using a single colony from a LB medium agar plate. The medium was supplemented with the appropriate antibiotics if required. Cultures were shaken at 200 rpm in culture tubes or baffled flasks at 37 °C. The incubation times varied depending on the desired optical densities.

II.4.4 Liquid cultures of *Bacillus megaterium*

Liquid starter-cultures in LB medium were inoculated using a single colony from a LB medium agar plate. The medium was supplemented with the appropriate antibiotics if required. Cultivations were performed at 37 °C in baffled flasks at 100 rpm in water bath shakers for approximately 14 h.

Liquid main cultures were inoculated at the ratio of 1:100 from the starter-cultures. Cultivations were performed at 37 °C in baffled flasks at 250 rpm in water bath shakers.

II.4.5 Determination of cell density

The cell densities of liquid cultures were determined by measuring the optical density (OD) at a wavelength of 578 nm. For cell densities with an OD_{578nm} > 0.8, appropriate dilutions of the cell culture broth were prepared. An OD_{578nm} of 0.5 corresponded to approximately 5 x 10^8 cells per mL. Biedendieck *et al.* (2007c) correlated, that an OD_{578nm} of 0.5 for *B. megaterium* MS941 equals 0.167 g cell dry weight (CDW).

II.4.6 Storage of bacteria

Strains were kept on LB medium agar plates at 4 °C for up to 10 days (*B. megaterium*) or up to 4 weeks (*E. coli*). For long-term storage of bacteria, glycerol cultures were prepared. For this purpose, liquid cultures of *B. megaterium* were incubated for approximately 14 h and 100 rpm at 37 °C in a water bath shaker. A culture volume of 650 µL was mixed with 350 µL of 87 % (w/v) glycerol. Stocks were immediately frozen and stored at -80 °C.

II.5 Molecular biology techniques

II.5.1 Preparation and transformation of chemically competent *Escherichia coli*

A volume of five milliliter of LB medium was inoculated with a single colony of *E. coli* DH10B, *E. coli* BL21 (DE3) or *E. coli* XL10-Gold and cultivated for approximately 14 h at 37 °C. One milliliter of this culture was used to inoculate 100 mL of LB medium. The bacteria were incubated aerobically at 37 °C in baffled flasks until the culture reached an OD_{578nm} of 0.6 - 0.8. After cooling on ice water for 10 min, the cells were harvested by centrifugation (4,500 x g; 10 min; 4 °C). The cells were suspended in 10 mL of ice-cold $CaCl_2$ solution and incubated on ice for 15 min. After centrifugation (4,500 x g; 10 min; 4 °C), the cells were suspended in 1 mL of ice-cold $CaCl_2$ solution. The competent cells were either used directly for transformation or were stored at -80 °C.

Transformation of chemically competent cells was applied as standard transformation procedure. Fifty microliter of $CaCl_2$-competent *E. coli* were mixed with either 1 µL of

plasmid solution (50 µg mL^{-1}) or 5 µL DNA of a ligation, incubated on ice for 20 min and subjected to a heat shock for 45 s at 42 °C. Immediately thereafter, the sample was cooled on ice for 2 min. Afterwards, 250 µL of LB medium were added and cells were incubated at 37 °C for 1 h. Depending on the expected transformation efficiency, different volumes were plated on LB medium agar plates containing the appropriate antibiotics. Incubation was performed overnight at 37 °C.

$CaCl_2$ solution

$CaCl_2$	100.0	mM
Glycerol	10.0	% (w/v)

II.5.2 Preparation and transformation of *Bacillus megaterium* protoplasts

Cells of an individual *B. megaterium* colony were grown in 50 mL of LB medium for approximately 14 h at 37 °C and 100 rpm in a water bath shaker in a baffled shaking flask. One milliliter of this culture was used to inoculate 50 mL of LB medium. The culture was incubated at 37 °C and 250 rpm in a water bath shaker in a baffled flask until it reached an OD_{578nm} of 1.0. Cells were sedimented by centrifugation (4,500 x g; 15 min; 4 °C) and suspended in 5 mL of SMMP solution. After addition of 100 µL of sterile lysozyme solution (100 µg of lysozyme solved in 1 mL SMMP solution), the protoplast suspension was incubated at 37 °C for 30 min under gentle shaking. Formation of protoplasts was monitored microscopically. After approximately 90 % of the rod shaped bacterial cells had formed into coccoid protoplasts, cells were harvested (1,400 x g; 10 min; RT). The supernatant was decanted and the protoplasts were carefully suspended in 5 mL of SMMP solution. After a second washing step, the protoplasts were suspended in 5 mL of SMMP solution and mixed with 750 µL of 87 % (w/v) glycerol. They were either used directly for transformation or were frozen and stored in aliquots of 500 µL at -80 °C for a period of up to 4 months.

Before transformation, protoplasts were tested for viability. Therefore, a 500 µL aliquot of protoplast solution was mixed with 3.5 mL of CR5-top agar and was plated

onto a LB medium agar plate without antibiotics. After incubation overnight, a thick film of *B. megaterium* cells should be seen.

For the transformation of protoplasts, approximately 1 µg of dried plasmid DNA was dissolved in 20 µL of SMMP solution at 37 °C for at least 20 min. Five hundred microliter of protoplast suspension were mixed with the DNA solution and transferred into tubes containing 1.5 mL of PEG-P solution. After incubation for 2 min at RT, 5 mL of SMMP solution were added and mixed carefully with the suspension. The protoplasts were sedimented by centrifugation (1,400 x g; 10 min; RT), suspended with caution in 500 µL of SMMP solution and incubated at 30 °C for 45 min in a thermomixer without shaking followed by 45 min with smooth shaking at 400 rpm. Subsequently, regenerated protoplasts were mixed with 3.5 mL of pre-warmed (42 °C) CR5-top agar and plated onto solid LB medium containing the required antibiotics. The plates were incubated at 30 °C for up to 24 h. Colonies arising after this period of incubation were plated onto new LB medium agar plates containing the required antibiotics.

SMMP solution

Antibiotic medium No. 3 (Difco)	17.5	g L^{-1}
Prepared as 2-fold stock solution and sterilized.		
Malic acid	20.0	mM
NaOH	40.0	mM
$MgCl_2$ x 6 H_2O	20.0	mM
Sucrose	500.0	mM
Prepared as 2-fold stock solution and sterilized by filtration (pore size 0.2 µm).		
Prior to use both stock solutions were mixed 1:1.		

PEG-P solution

PEG 6000	400.0	g L^{-1}
Malic acid	20.0	mM
NaOH	40.0	mM
$MgCl_2$ x 6 H_2O	20.0	mM
Sucrose	500.0	mM
Sterilized by autoclaving (11 min, 120 °C).		

CR5-top agar (3.5 mL)

Solution A	1750.0	µL
Solution B	998.0	µL
8 x CR5 salts	346.0	µL
L-proline (12 % (w/v))	175.0	µL
D-glucose (20 % (w/v))	175.0	µL

Solution A

Sucrose	602.0	M
MOPS	58.0	mM
NaOH	30.0	mM

The pH was adjusted to 7.3 using NaOH. Sterilization was done by sterile filtration (0.2 µm pore size).

Solution B

Agar-agar	1.40	% (w/v)
Casamino acids	0.07	% (w/v)
Yeast extract	3.51	% (w/v)

Sterilized by autoclaving.

CR5-Salts

K_2SO_4	11.0	mM
$MgCl_2$ x 6 H_2O	394.0	mM
KH_2PO_4	3.0	mM
$CaCl_2$	159.0	mM

Sterilized by autoclaving.

II.5.3 Preparation of plasmid DNA from *Escherichia coli*

Plasmid DNA destined for protoplast transformation of *B. megaterium* was prepared from *E. coli* DH10B or XL10-Gold carrying the corresponding plasmid using the QIAprep miniprep kit according to the manufacturer's instructions. Alternatively, 5 mL of an overnight culture were harvested by centrifugation (15,000 x g; 2 min; RT) and the cells were suspended in 300 µL of buffer P1. Three hundred microliter of buffer

P2 were added, the sample was mixed by inverting the tube and incubated at RT for 2 min. Subsequently, 300 µL of buffer P3 were added and the sample was mixed again. After a centrifugation step (15,000 x g; 30 min; RT) 900 µL of the supernatant were added to 633 µL isopropanol in a new tube. Precipitation and sedimentation of plasmid DNA took place during 30 min centrifugation at RT (15,000 x g). The DNA was washed with 400 µL of 70 % (v/v) ethanol, centrifuged again (15,000 x g; 5 min; RT) and dried for 15 min at 70 °C in a thermomixer. After all traces of ethanol were evaporated, the DNA was dissolved in 50 µL dH_2O.

Buffer P1

Tris-HCl (pH 8.0)	50.0	mM
EDTA	10.0	mM
RNase A	100.0	µg mL^{-1}

Buffer P2

NaOH	200.0	mM
Na-dodecyl-sulfate	1.0	% (w/v)

Buffer P3

K-acetate	3.0	M

The pH was adjusted to 5.5 using acetic acid.

II.5.4 Quantification of nucleic acids

The nucleic acid concentration of DNA and RNA samples was determined using the Nanodrop ND-1000 spectrophotometer. For quantification of nucleic acids the absorbance at 260 nm was measured. For pure nucleic acid solution an OD_{260nm} of 1 corresponds to a double-stranded DNA concentration of 50 µg mL^{-1} or a single-stranded RNA concentration of 40 µg mL^{-1}. Protein impurities of the nucleic acid solution were deduced from the ratio of OD_{260nm} to OD_{280nm}. Nucleic acid samples having an OD_{260nm} to OD_{280nm} ratio of 1.8 - 2.0 were considered as pure.

II.5.5 Electrophoretic separation of DNA

For the analytical separation of DNA-fragments agarose gels (0.5 - 2 % (w/v) agarose in TAE-buffer) were prepared. DNA samples were mixed with DNA loading dye to facilitate loading and to indicate the progress of DNA migration. GeneRuler™ DNA ladder mix or MassRuler DNA ladder mix were used as size standards according to the manufacturer's instructions. For the electrophoretic separation a voltage of 100 V was applied. Afterwards, gels were incubated in an ethidium bromide solution for 15 min. Subsequently, the DNA was visualized by the fluorescence of ethidium bromide under UV light (λ = 312 nm).

TAE-buffer

TRIS-acetate (pH 8.0)	40.0	mM
EDTA	1.0	mM

DNA loading dye

Bromphenol blue	350.0	µM
Cylene cyanol FF	450.0	µM
Glycerol	50.0	% (w/v)

Ethidium bromide solution

Ethidium bromide (1.0 % stock solution)	0.1	% (v/v)

II.5.6 Amplification of DNA fragments by polymerase chain reaction (PCR)

For amplification of DNA by PCR, oligonucleotide primers for each DNA fragment of interest were designed. If appropriate, recognition sequences for restriction endonucleases were inserted in the primer sequence. Site-directed mutagenesis PCRs were performed with the Quik Change II kit (Agilent) according to the manufacturer's instructions. All oligonucleotide primers used in this work are listed in Table II.f. Primers were purchased from Biomers or Metabion.

Typically, 10 ng genomic DNA or 1 ng plasmid DNA were used as templates in PCR reactions. Phusion™ polymerase (Finnzymes) was applied for error-proof DNA applification purposes. Taq DNA polymerase (New England BioLabs) was used in all other cases. Both polymerases were applied according to the manufacturer's instructions.

Table II.f: Oligonucleotide primers used in PCR reactions. Nucleotides comprising important endonuclease recognition sequences are printed in italics. Base exchanges in primers for site-directed mutagenesis are indicated with capitals.

Primers	Application	Sequence (5'-3')
QC-p3STOP1622-for	generation of p3STOP1622	ccaagaattggagcTaatTaattcttgcggag
QC-p3STOP1622-rev		ctccgcaagaattAattAgctccaattcttgg
QC-p3STOP1623-for	generation of p3STOP1623	ctaaaaaaaacattgaaataaacat*ttaAttAA*tatatg atgagataaagttagtttattgg
QC-p3STOP1623-rev	generation of p3STOP1623	ccaataaactaactttatctcatcatata*TTaaTtaa*at gtttatttcaatgttttttttag
gfp-for	amplification of *gfp*	ttatt*agatct*gggctagcaaaggagaagaac
gfp-rev		atat*cggccg*cgaattcattatttgtagagc
QC-35^{+}-for	generation of pSSBm74 and pSSBm84	ctaaattttctaaaaaaaacattgaCataaacatttaatt aatatatgatgagataaag
QC-35^{+}-rev		ctttatctcatcatatattaattaaatgtttatGtcaatgttttttt tagaaaatttag
QC-Δ*Nhe*I-for	generation of pSSBm85	gataaacaaactaactcaattaagAtagTtgatggata aacttgttcac
QC-Δ*Nhe*I-rev		gtgaacaagtttatccatcaActaTcttaattgagttagtt tgtttatc
sp$_{yngK}$-for	generation of pSSBm22	gcag*tgtaca*atgtatattaaaaaatgtatagg
sp$_{yngK}$-rev		tata*agatctactagt*atctgctttagcaggtaacg
sp$_{vpr}$-for	generation of pSSBm23	tagt*tgtaca*atgaaaaaagtatcaattcg
sp$_{vpr}$-rev		tata*agatctactagt*gttggcggatgctcccattgc
sp$_{nprM}$-for	generation of pSSBm24	ttat*tgtaca*atgaaaaagaaaaaacagg
sp$_{nprM}$-rev		tata*agatctactagt*cgcggcactgctcgtatgtgc
sp$_{yocH}$-for	generation of pSSBm25	tata*tgtaca*atgaaaaagacaatgattacg
sp$_{yocH}$-rev		tata*agatctactagt*tgcagctgatgcaccacttgc
sp$_{asp}$-rev	amplification of *sp$_{asp}$*	tata*actagt*atcagcagatgctgttgttgg

Primers	Application	Sequence (5'-3')
sp_{pac}-rev	amplification of sp_{pac}	tata*actagt*cccagcaaaaactagattttg
$sp_{asp/pac}$-for	amplification of sp_{asp} and sp_{pac}	ccgactagaccataaggg
coco-for	amplification of "on command" plasmid copy number control elements	aatat*gcatgc*taactagctaaccaattctcatgtttgac agc
coco-rev		aatat*gacgtc*ggccgcatcgaatataac
tfh-for	amplification of *tfh-his*$_6$	ttta*actagt*gctaacccatatgaacgtgg
tfh-rev		tata*cggccg*ttagtgatggtg
trfA-for	amplification of *trfA*	ccataagattagc*ggatcc*tacc
trfA-rev		atata*gcatgc*ttaaacgcctggttgc
QC-*Spe*I-for	generation of pBmEcoco	catctgcatcaagaacAgtttaagctcacgac
QC-*Spe*I-rev		gtcgtgagcttaaactTttcttgatgcagatg
QC-*Xho*I-for	generation of pBmEcoco	ggcgaagccctATatcccctgcc
QC-*Xho*I-rev		ggcaggggatctATggcttcgcc
K1EP_a-for	amplification of *k1ep*	atata*gaattc*caagacctacacgctattcaac
K1EP_a-rev		atata*cggccg*ttatgcgaaacaatagtctg
pSSEc4-for	generation of pSSEc4	atatt*gaattc*tatagtgtccagtaaatgtgaaattgttatc cgctcac
pSSEc4-rev		gcgaggaagcggaagagcg
NIAysic1-for	amplification of the insert of pNIAysic1	atatt*gaattc*gcag*gtcgac*ggtatgg
NIAysic1-rev		aatata*agctt*ttttttttttttttttttttttttttttttttgggagctcgc
K1EP_b-for	amplification of *k1ep*	atata*ctagt*caagacctacacgctattcaac
K1EP_b-rev		tata*gagctc*ttatgcgaaacaatagtctg
SP6P-for	amplification of *sp6p*	atat*gagctc*aagatttacacgctatccag
SP6P-rev		tatat*ggatcc*ttaggcaaatacgtattcag

II.5.7 Phosphorylation and hybridization of oligo-nucleotides

Three hundred picomol of oligonucleotides were phosphorylated using 10 U of T4 polynucleotide kinase (New England BioLabs) according to the manufacturer's instructions. After heat inactivation of the enzyme at 65 ℃ for 20 min the two

corresponding phosphorylated oligonucleotides were mixed and incubated at 95 ℃ for 3 min. Annealing was allowed to proceed at the calculated T_m-2 ℃ for 1 min. After subsequent incubation at 55 ℃ for 1 min the hybrid ized oligonucleotides were used in ligation reactions.

The oligonucleotides used in the study are listed in Table II.g.

Table II.g: Oligonucleotides. Nucleotides belonging to important endonuclease recognition sequences are printed in italics. Base exchanges in oligonucleotides replacing their plasmid DNA pendants are indicated in bold.

Oligonucleotide	Application	Sequence (5'-3')
p3STOP1624-for	generation of p3STOP1624	*taa*tatatgatgagataaagttagtttattggataaacaaactaact caattaa*gCtaGC*tgatggataaacttgttcacttaaatcaaag ggggaaa*t*
p3STOP1624-rev		*gtaca*tttccccctttgatttaagtgaacaagtttatccatca*GCta Gc*ttaattgagttagtttgtttatccaataaactaactttatctcatcat ata*ttaat*
-10$^+$-for	generation of pKMBm1	*taa*tatatAatgagataaagttagtttattggataaacaaactaac tcaattaa*g*
-10$^+$-rev		*ctagc*ttaattgagttagtttgtttatccaataaactaactttatctcat Tatata*ttaat*
utr$^+$-for	generation of pSSBm40 and pKMBm4	*taa*tatatgatggaattgtagttagtttacaattccaacaaactaac tcaattaa*g*
utr$^+$-rev		*ctagc*ttaattgagttagtttgttggaattgtaaactaactacaattcc atcatata*ttaat*
rbs$^+$-for	generation of pSSBm44	*ctagc*tgatggataaacttgttcacttaaatcaaggAggTGa*at*
rbs$^+$-rev		*gtaca*ttCAccTccttgatttaagtgaacaagtttatccatca*g*
sp$^+_{yngK}$-for	generation of pSSBm27	*gtacaa*tgtatattaaaaaatgtattggttctattttattcttattattatt ctgttcttctgctttaccagctaaagctgat*actagta*
sp$^+_{yngK}$-rev		*gatctactagt*atcagctttagctggtaaagcagaagaacagaa taataataagaataaaatagaaccaatacatttttaatatacat*t*
pK1Eivt-for	generation of pK1Eivt	*ccggg*taccgagctcgGttcgtaatcatGtcatagctgtttcctgt gtgaaatt*g*
pK1Eivt-rev		*aattc*aatttcacacaggaaacagctatgaCatgattacgaatC gagctcggta*c*
P_{K1E-1}-for	generation of pP_{K1E-1}	*ttaag*atcagtatttactggacactatagaagggagaaatgtctcc ctaaattatcaatttgatttataaggaggtgaa*at*
P_{K1E-1}-rev		*gtaca*tttcacctccttataaatcaaattgataatttagggagacatt tctcccttctatagtgtccagtaaatactgat*c*

Oligonucleotide	Application	Sequence (5'-3')
P_{K1E-2}-for	generation of pP_{K1E-2}	*ttaag*caaactatttagctgacactataagagaaggcttaacaag gcgttgctacggtagcgcctgattaaactttcacttacaaggaggt gaaa*t*
P_{K1E-2}-rev		*gtaca*tttcacctccttgtaagtgaaagtttaatcaggcgctaccgt agcaacgccttgttaagccttctcttatagtgtcagctaaatagtttg *c*
P_{SP6}-for	generation of pP_{SP6}	*ttaag*catacgatttaggtgacactatagaatagaagtatagtgc cgttcttttgagcggcctattactcaccagtcttcacggggagggct ggatagtaaaaggaggtgaaa*t*
P_{SP6}-rev		*gtaca*tttcacctccttttactatccagccctccccgtgaagactggt gagtaataggccgctcaaaagaacggcactatacttctattctat agtgtcacctaaatcgtatg*c*
term-for	generation of $pP_{K1E-1+t}$, $pP_{K1E-2+t}$, pP_{SP6+t}	*c*taactagcta*accggt*tacctatgccctatctaccttgcgtaggta gggttctttttgtt*g*
term-rev		*ccggc*aacaaaaagaaccctacctacgcaaggtagatagggc ataggta*accggt*tagctagtta*gcatg*

II.5.8 Enzymatic digestion of DNA

Site-directed digestion of double stranded DNA (vectors and PCR products) was carried out using restriction endonucleases (New England BioLabs, Fermentas). Reaction buffers, concentrations of enzymes and DNA concentrations as well as incubation temperatures were chosen according to manufacturer's instructions. The restrictions were allowed to proceed for more than 30 min or overnight at 37 °C. Subsequently, gel electrophoretic separations of the resulting fragments were performed (section II.5.5). For restriction analysis the fragments were stained with ethidium bromide, while DNA for cloning purposes was visualized using the GelStar Nucleic Acid Gel Stain (Lonza), a blue light source (Flu-O-blu, Biozym) and a yellow filter. The DNA fragments of interest were excised from the gels and purified using the QIAquick gel extraction kit (Qiagen) according to manufacturer's instructions.

II.5.9 Ligation of DNA fragments

Twenty five to two hundred nanogram of plasmid DNA were applied in ligation reactions. Insert-DNA was added in excess (ratio with regard to molar concentrations of 2:1 to 10:1). Ligations were carried out in 10 µL batches according to the

manufacturer's instructions using T4 DNA ligase (New England BioLabs). All reactions were incubated for at least 30 min or overnight at 25 °C.

II.5.10 DNA sequencing

The successful modification of DNA was confirmed by sequence determination of the respective DNA region based on the principle of the Sanger dideoxy-method (Sanger *et al.*, 1977). The sequencing reactions were either conducted on site with subsequent analysis with the ABI PRISM 310 Genetic Analyzer (Applied Biosystems) or the analysis was commissioned off-site (GATC Biotech AG). If sequencing was performed on-site the required preparatory PCR with fluorescence-labelled ddNTPs and the purification of the PCR product were carried out as described by the manufacturer. The analysis of all sequencing results was done using the computer software Sequence Analysis v5.2 (Applied Biosystems), Chromas (Technelysium) and Vector NTI (Invitrogen).

II.6 Protein biochemical techniques

II.6.1 Expression of recombinant genes in *Bacillus megaterium*

The respective *B. megaterium* plasmid strains were streaked onto LB medium agar plates containing the appropriate antibiotics and were cultivated for approximately 14 h at 37 °C. Single colonies were used to inoculate 50 mL of LB medium supplemented with appropriate antibiotics. The cultivation was performed at 37 °C for approximately 14 h with gentle shaking at 100 rpm in a water bath shaker in baffled flasks. One milliliter of this starter-culture was used to inoculate 100 mL of LB growth medium supplemented with appropriate antibiotics. Cultivations were performed in baffled flasks at 37 °C and shaking at 250 rpm in water bath shakers. After reaching an OD_{578nm} of 0.3 - 0.4, recombinant gene expression was induced by the addition of 0.5 % (w/v) xylose.

II.6.2 Analysis of basal *gfp* expression in *Escherichia coli*

E. coli DH10B strains harboring the corresponding plasmid constructs were plated on solid LB medium with appropriate antibiotics and cultivated over night at 37 °C. Starter-cultures of 50 mL LB medium with appropriate antibiotics and indicated additives were inoculated from single colonies and were cultivated at 37 °C and 250 rpm in a water bath shaker for 14 h. Hundred milliliter LB main cultures with appropriate additives were inoculated 1:50 from the starter-culture and cultivated at 37 °C in a water bath shaker with 250 rpm. Gfp fluo rescence was measured after 8 h of cultivation (see section II.6.9).

II.6.3 Production of recombinant proteins in *Escherichia coli*

E. coli BL21(DE3) cells were transformed with the plasmid pET-32a(+)-K1EP and employed for production of the His_6-tagged K1E phage originated RNAP. Baffled shaking flasks with LB medium were inoculated with a final concentration of 5 x 10^7 cells per mL of a starter-culture and subsequently cultivated at 37 °C with vigorous shaking. At an OD_{578nm} of 0.5, Isopropyl-β-D-thiogalactopyranosid (ITPG) was added to the growth medium to a final concentration of 100 µM. Approximately 4.5 h after induction of heterologous gene expression cells were sedimented at 4 °C and the target protein was purified after cell disruption (see section II.6.7).

II.6.4 Preparation of intracellular protein fractions

Approximately 3 x 10^9 *B. megaterium* cells were harvested by centrifugation (15,000 x g; 10 min; 4 °C). The supernatant was rem oved completely and the cells were suspended in 30 µL of lysis buffer. After subsequent incubation at 37 °C for 30 min with vigorous shaking at 1,000 rpm in a thermomixer, the samples were centrifuged (15,000 x g; 30 min; 4 °C). Twenty six microliter of the supernatant containing the soluble proteins were mixed with 13 µL of SDS loading dye. The sediment, containing cell debris and insoluble proteins, was suspended in 30 µL of urea buffer and centrifuged again (15,000 x g; 30 min; 4 °C). Twenty-six microliter of the supernatant were mixed with 13 µL of SDS loading dye. Seven and a half

microliter of the prepared samples which corresponded to proteins equivalents of 0.5 x 10^9 cells were analyzed by SDS-PAGE (section II.6.6).

Lysis buffer

Na_3PO_4	100.0	mM
$MgSO_4$	2.0	mM
Lysozyme	0.5	% (w/v)
Benzonase	500.0	U mL^{-1}

The pH was adjusted to 6.5 using H_3PO_4. Aliquots of lysis buffer were stored at -20 °C for several months. Benzonase (Merck KGaA) was freshly added prior to use.

Urea buffer

Tris-HCl (pH 7.5)	50.0	mM
Urea	8.0	M

SDS loading dye

Tris-HCl (pH 6.8)	100.0	mM
Glycerol	40.0	% (w/v)
β–mercaptoethanol	2.0	mM
SDS	110.0	mM
Bromophenol blue	3.0	mM

II.6.5 Preparation of extracellular protein fractions

Extracellular proteins were precipitated by ammonium sulfate. For this purpose, cell-free supernatant samples were obtained by centrifugation (4,500 x g; 15 min; 4 °C). 1.5 mL of the supernatant were mixed with ammonium sulfate to a final concentration of 70 % (w/v). The samples were incubated with gentle mixing at 4 °C for 2 h. Precipitated proteins were collected by centrifugation (15,000 x g; 30 min; 4 °C). After removal of the supernatant, the proteins were suspended in 10 µL of urea buffer (see section II.6.4). Five microliter of SDS loading dye (see section II.6.4) were added and the samples were analyzed by SDS-PAGE (section II.6.6).

II.6.6 Electrophoretic separation of proteins (SDS-PAGE)

Proteins were analyzed by SDS-polyaclyamide gel electrophoresis (PAGE) as described by Laemmli (1970) with modifications by Righetti (1990) for discontinuous SDS-PAGE. Protein samples were denatured at 95 °C for 5 min in SDS loading dye (see section II.6.4). Samples were loaded onto the gel which was run approximately 40 min at 45 mA. During electrophoresis, proteins were first focussed in the stacking gel and subsequently separated according to their relative molecular mass in the running gel. The size standards employed were protein molecular weight marker, PageRuler prestained protein ladder or PageRuler unstained protein ladder (Fermentas). Gels were stained (staining solution) for at least 1 h and subsequently destained (destaining solution) until the protein bands were clearly visible.

Stacking gel (6 % (v/v)), sufficient for 4 mini-gels

dH_2O	5.5	mL
Tris-HCl (500 mM, pH 6.8); 0.4 % (w/v) SDS	2.5	mL
Acrylamide stock solution (30 %)	2.0	mL
TEMED	10.0	µL
Ammonium peroxodisulfate solution (10 % (w/v))	100.0	µL

Running gel (12 % (v/v)), sufficient for 4 mini-gels

dH_2O	7.0	mL
Tris-HCl (1.5 M, pH 8.8); 0.4 % (w/v) SDS	5.0	mL
Acrylamide stock solution (30 %)	8.0	mL
TEMED	20.0	µL
Ammonium peroxodisulfate solution (10 % (w/v))	200.0	µL

Electrophoresis buffer

Tris-HCl (pH 8.8)	50.0	mM
Glycine	385.0	mM
SDS	0.1	% (w/v)

Staining solution

Ethanol	30.0	% (v/v)
Acetic acid	10.0	% (v/v)
Coomassie brilliant blue G-250	0.3	% (w/v)

Destaining solution

Ethanol	30.0	% (v/v)
Acetic acid	10.0	% (v/v)

II.6.7 Purification of His_6-tagged fusion proteins

E. coli host cells harboring His_6-tagged target proteins were washed with lysis buffer. Subsequently, bacteria suspended in lysis buffer were disrupted by a double passage through a french pressure cell press (Thermo) at 19,200 psi. Subsequently cell debris was sedimented by ultra-centrifugation (110,000 x g; 4 ℃) and the supernatant was subjected to affinity chromatography using a Ni-NTA column (Qiagen) equilibrated with lysis buffer. After a washing step with buffer A, pre-elution was performed with buffer A containing 20 mM imidazole. For the final elution of His_6-tagged K1E-RNAP, buffer A containing 250 mM imidazole was used. Subsequently, the RNAP was dialyzed at 4 ℃ against buffer B using a 10.000 MWCO membrane. After dialysis 50 % glycerol were added for long-term storage at -20 ℃.

Lysis buffer

$NaPO_4$	20.0	mM
NaCl	100.0	mM

The pH was adjusted to 7.7.

Buffer A

$NaPO_4$	20.0	mM
NaCl	500.0	mM

The pH was adjusted to 7.7.

Buffer B

$NaPO_4$	20.0	mM
NaCl	100.0	mM
Ethylenediaminetetraacetic acid (EDTA)	1.0	mM
1,4-Dithiothreitol (DTT)	10.0	mM
Triton X-100	0.1	% (v/v)

The pH was adjusted to 7.7.

II.6.8 Quantification of purified protein

For quantification of purified protein, a Bradford protein assay kit (Biorad) was applied according to the manufacture's instructions using bovine serum albumin (BSA) (Sigma-Aldrich) as standard.

II.6.9 Fluorescence measurements of green fluorescent protein

Recombinant Gfp was quantified via fluorescence spectroscopy measurements using the Luminescence Spectrometer LS50B (PerkinElmer) and a fluorescence cuvette (type 104.002F-QS, Helma). For this purpose, bacterial cells were harvested by centrifugation (15,000 x g; 10 min; 4 °C). The resu lting cell sediment was suspended in 100 mM sodium phosphate buffer (pH 7.0) to a final concentration of 1 x 10^9 cells per mL. For *B. megaterium* Gfp productions, 100 µL of these samples were diluted with 900 µL buffer and mixed well. The cell samples were excited using a wavelength of 475 nm, while fluorescence emission was recorded at 512 nm. The relative fluorescence of different amounts of purified Gfp was determined and quantified using the following linear correlation: Gfp [mg mL^{-1}] = relative emission maxima x 3.42 x 10^{-6} x dilution factor. Since Biedendieck *et al.* (2007c) correlated, that 0.334 g L^{-1} cell dry weight (CDW) equal an OD_{578nm} of 1 for *B. megaterium* MS941, the Gfp amount could also be specified in mg g_{CDW}^{-1}. Here, this equation originally obtained for purified Gfp was applied to estimate Gfp amounts inside bacterial cells. Due to additional light absorption and scattering by the cells a slight underestimation of the produced Gfp has to be taken into account.

II.6.10 Determination of hydrolase activity

Hydrolase activity of the Tfh enzyme was measured by detection of the hydrolysis product of *p*-nitrophenylpalmitate (*p*NPP). For this purpose, 40 µL of cell free culture supernatant were added to 960 µL of freshly prepared *p*NPP solution and mixed well. The enzyme activity assay was carried out at 30 °C. The enzymatic release of *p*-nitrophenol was photometrically detected at 400 nm for 90 s using the spectrophotometer V-550 (Jasco). One enzyme unit [U] was defined as the amount that caused the release of 1 µmol *p*-nitrophenol per minute under the given assay conditions.

$$EA\left[U \cdot mL^{-1}\right] = \frac{\Delta A\left[min^{-1}\right] \cdot V_R\left[mL\right]}{\varepsilon\left[cm^2 \cdot \mu mol^{-1}\right] \cdot d\left[cm\right] \cdot V_E\left[mL\right]}$$

Formula II.a: Formula for the calculation of the volumetric hydrolase activity.

EA = volumetric enzyme activity

ΔA = change in absorbance

V_R = reaction volume

V_E = volume of enzyme solution

ε = extinction coefficient of *p*-nitrophenol (9.62 cm^2 μmol^{-1})

d = thickness

For the calculation of hydrolase amounts [mg L^{-1}], the specific activity value for purified enzyme determined by Yang *et al.* (2007) and a temperature correlation factor were employed.

Solution 1

p-nitrophenylpalmitate	0.30	% (w/v)
Dissolved in isopropanol.		

Solution 2

Taurocholic acid sodium salt hydrate	0.23	% (w/v)
Gum arabic	0.11	% (w/v)
Dissolved in Na-phosphate buffer (20 mM, pH 7.1).		

***p*NPP solution**

Mixture of Solution 1 and Solution 2 at a ratio of 1 : 10.

II.6.11 Phage RNA polymerase driven *in vitro* transcription

RNAP activity was determined by quantification of *in vitro* generated transcription products. Formed RNA chains were quantified by capillary electrophoresis using the RNA Pico 6000 kit on the 2100 Bioanalyzer platform (Agilent).
Conditions for the first *in vitro* transcriptions were appropriate amounts of K1E RNAP in the initial reaction buffer and 10 ng μL^{-1} plasmid DNA as template.

Initial reaction buffer

Tris-HCl (pH 7.9)	40.0	mM
$MgCl_2$	6.0	mM
Spermidine	2.0	mM
NaCl	10.0	mM
1,4-Diothiothreitol (DTT)	10.0	mM
RNAse inhibitor (RNAse OUT, Invitrogen)	2.0	U μL^{-1}
of each rNTP	0.5	mM

During the systematic optimization approach the indicated parameters were changed to estimate the optimum conditions for K1E RNA polymerase activity *in vitro*. Enzymatic reactions were performed at 37 ℃ for 30 minutes with gentle shaking. The enzyme reaction was stopped by addition of 23 mM EDTA to the reaction mixture and additionally placing on ice.
One unit was defined as the enzyme amount capable to incorporate 1 nmol rNTPs into RNA during a 30 minute reaction at 37 ℃ in 10 µL under optimized reaction buffer conditions with 10 ng μL^{-1} $pP_{K1E\text{-}1+t}$-*gfp* as template.

Optimized reaction buffer

Tris-HCl (pH 7.9)	40.0	mM
$MgCl_2$	10.0	mM
Spermidine	1.3	mM
NaCl	10.0	mM
1,4-Diothiothreitol (DTT)	10.0	mM
RNAse inhibitor (RNAse OUT, Invitrogen)	2.0	U μL^{-1}
of each rNTP	0.5	mM

Large scale *in vitro* transcriptions were performed using 27 ng μL^{-1} of linearized pK1Eivt-NIAysic1 as template, 10 U of K1E RNAP and optimized transcription buffer conditions. The enzymatic reactions were done in 50 µL assays at 37 ℃ for 2.5 h. RNA synthesis was followed by DNA template removal. For this purpose, the reaction mixture was incubated with 1.5 U of RQ DNase (Promega) for 30 minutes. The subsequent RNA purification using the innuPREP RNA Mini Kit (Analytik Jena) was performed according to the manufacturer's instructions.

II.7 Fluorescence microscopy

Fluorescence microscopy was performed with several *B. megaterium* plasmid strains. Cells were diluted appropriately in medium and were analyzed microscopically using the Axiovert 200M (Carl Zeiss AG). Each image section was documented under phase contrast and fluorescence excitation using appropriate filters for the detection of Gfp.

III. Results and Discussion

III.1 Directed optimization of the P_{xylA}-based gene expression system

The major goal of this part of the work was the significant enhancement of the recombinant intra- and extracellular protein production with *B. megaterium*. Based on the previously established plasmid system for xylose-inducible recombinant protein production (Malten *et al.*, 2006; Rygus & Hillen, 1991) various directed genetic optimization strategies were followed.

The quantity of an intracellular protein formed by a microbial cell is mainly determined by several closely interconnected steps: transcription, mRNA stability and translation. Consequently, promoter sequences, mRNA signatures and ribosome binding site of the plasmid-borne P_{xylA}-controlled protein production system were target for directed genetic optimization approaches.

III.1.1 Rationale of the approach and establishment of the basal production vector

A first optimization of the *B. megaterium* expression plasmid pSTOP1622 was done by insertion of two additional stop codons downstream of the *mcs* creating the plasmid p3STOP1622. This new plasmid enables the cloning of genes lacking a stop codon without risking the generation of mRNAs encoding translational fusions with vector parts. To make this new basic production plasmid p3STOP1622 susceptible to multiple genetic optimizations, two new unique restriction sites were introduced between the promoter P_{xylR} and the translational start site located upstream of the multiple cloning site to allow for a cassette exchange strategy. Since DNA changes in this region might affect the production of a target protein, both sites were chosen to cause only minimal changes in the native DNA sequence. By introduction of a *Pac*I restriction site between the -35 and -10 region of the promoter P_{xylA} the plasmid p3STOP1623 was created. Subsequently, an additional *Nhe*I recognition site was inserted downstream of the P_{xylA} transcriptional start site into the plasmid

p3STOP1623 resulting in p3STOP1624. The effects of these DNA modifications on protein production in *B. megaterium* were evaluated. For this purpose, these plasmids and the starting plasmid p3STOP1622 were equipped with *gfp* as model gene. Gfp is an excellent reporter protein, since it is not harmful to cells and is independent of substrates and cofactors, readily measurable and provides high sensitivity (Reischer *et al.*, 2004; Scholz *et al.*, 2000). *B. megaterium* MS941 cells were transformed with each of these plasmids (p3STOP1622-*gfp*, p3STOP1623-*gfp* p3STOP1624-*gfp*) individually and Gfp production experiments were carried out.
These showed that both restriction sites have some impact on target protein formation by *B. megaterium*. The introduction of the *Pac*I recognition site caused an increase in Gfp production of approximately 50 % from 4.6 mg g_{CDW}^{-1} to 6.8 mg g_{CDW}^{-1}, while the additional insertion of the *Nhe*I site reduced the Gfp production by 40 % back to 4.2 mg g_{CDW}^{-1}. In summary only minor changes were observed. Consequently, p3STOP1624 harboring both new restriction sites was used as starting point for further genetic optimizations.

III.1.2 Genetic elements enhancing heterologous protein production in *Bacillus megaterium*

Based on p3STOP1624 (Figure III.a) different stages of the protein formation process were target of genetic optimizations. The -35 (TTGAAA) and the -10 region (TATGAT) of P_{xylA} were changed into their predicted optimal counterparts termed -35^{+} (TTGA<u>C</u>A) and -10^{+} (TAT<u>A</u>AT), respectively. This usually leads to increased RNA polymerase binding, a higher transcription initiation frequency and thereby to an acceleration of the whole transcriptional process (deHaseth *et al.*, 1998; Qiu & Helmann, 1999). The DNA region surrounding the palindromic XylR-binding motif (in italics) termed *utr* (*AGTTAGTTT*ATTGGATAAAC*AAACTAACT*) was modified (underlined letters) to utr^{+} (<u>GGAATTGT</u>*AGTTAGTTT*<u>ACAATTCCAAC</u>*AAACTAACT*) for an increased mRNA half-life time via formation of a hairpin loop in the 5' untranslated RNA region of the target protein's mRNA. An optimized ribosome-binding site (Shine Dalgarno sequence) for *B. megaterium* (AAGG<u>A</u>GG<u>TG</u>A; modified bases are underlined) was already developed and successfully employed for recombinant production of the dextran sucrase DsrS (Malten *et al.*, 2005). This

adapted ribosome-binding site (*rbs*$^+$) is highly complementary to the 3'-end of 16SrRNA of the *B. megaterium* ribosome. It increases the ribosome's affinity to the mRNA of interest and thus enhances the translation process.

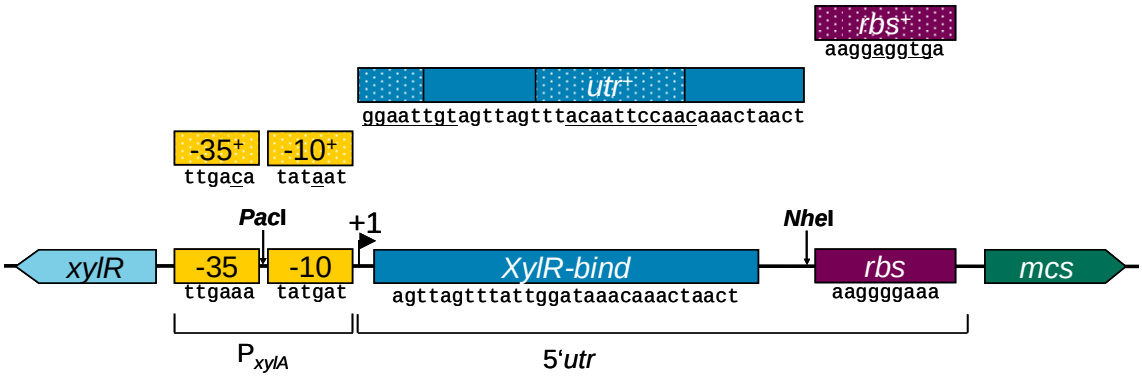

Figure III.a: Elements for regulated gene expression of the modified basic expression plasmid p3STOP1624.

Elements for gene expression in *B. megaterium* are the xylose-inducible core promoter (P_{xylA}) consisting of -35 and -10 region, the gene for the XylR repressor and its DNA binding region (*XylR-bind.*). Addition of xylose leads to conformational changes of XylR resulting in the release from the XylR-binding region and the initiation of transcription. The multiple cloning site (*mcs*) for introduction of target genes is located downstream of the native ribosome-binding site (*rbs*). The unique restrictions sites introduced in this work are indicated. Above the plasmid the optimized genetic elements (-35$^+$/-10$^+$/*utr*$^+$/*rbs*$^+$), which were evaluated in this work, are presented. Base exchanges or introductions are underlined.

After exchanging each of the mentioned genetic elements of p3STOP1624 into their improved counterparts individually, the resulting plasmids were equipped with *gfp* as reporter gene. *B. megaterium* MS941 cells were transformed with the different plasmids and protein production experiments were carried out. Cells harboring the basic plasmid p3STOP1624-*gfp* were used as reference. When tested separately all improved elements showed the expected enhancing effect on Gfp production in *B. megaterium* (Figure III.b). The Gfp amounts visualized by SDS-PAGE correlated well with the measured Gfp fluorescence values (compare Figure III.b, panel (a) with panel (b)). Obviously, Gfp was properly folded and active. Consequently, the overall Gfp production could reliably be quantified by fluorescence.

B. megaterium cells carrying pKMBm9 (-10$^+$) contained a nearly 3-fold higher Gfp amount (20.4 mg g_{CDW}^{-1}) than the reference cells (7.0 mg g_{CDW}^{-1}). The optimized

-35 promoter sequence (-35$^+$) (pSSBm84) revealed the highest positive impact on protein formation. Compared to the reference, the Gfp quantity was increased 11-fold to 76.5 mg g_{CDW}^{-1}. The modified element causing a hairpin loop in the 5' *utr* of the target mRNA (*utr*$^+$) (pSSBm46) enhanced Gfp production 3.8-fold (26.6 mg g_{CDW}^{-1}). The positive impact of an optimized ribosome-binding site on protein formation originally described by Malten *et al.* (2005) was clearly confirmed for the production of Gfp. *B. megaterium* cells harboring pSSBm50 encoding this optimized Shine Dalgarno sequence (*rbs*$^+$) contained double the amount of Gfp (14.8 mg g_{CDW}^{-1}) compared to the reference cells.

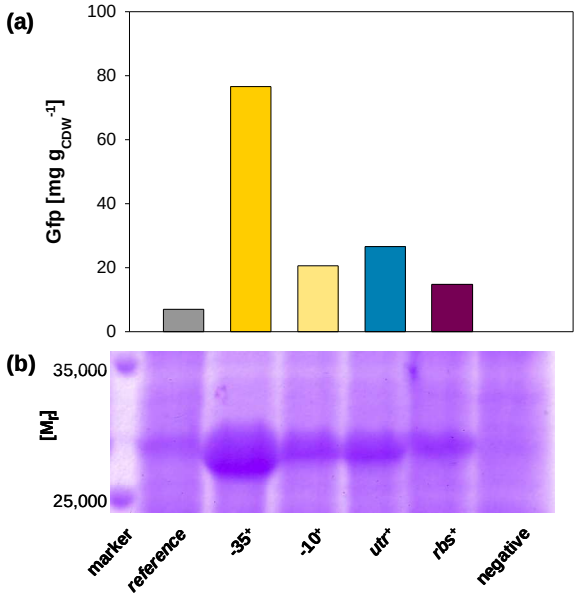

Figure III.b: Improved model protein production by optimized expression plasmids.
B. megaterium MS941 cells were transformed with one of the optimized expression plasmids carrying *gfp* under the control of the xylose-inducible promoter. Employed genetically optimized parts (-35$^+$ region, -10$^+$ region, DNA sequence encoding the modified 5' *utr*$^+$ and the optimized ribosome binding site (*rbs*$^+$)) of the expression plasmids are indicated. Cells harboring the non-optimized parental expression vector p3STOP1624-*gfp* were used as reference. A cell sample of MS941 p3STOP1624-*gfp* culture without addition of xylose served as negative control (negative). All cultivations were performed in LB medium supplemented with appropriate antibiotics at 37 °C. Heterologous production of Gfp was induced with 0.5 % (w/v) xylose at OD_{578nm} = 0.4. Samples were taken 7.5 h after xylose addition. **(a)** Gfp amounts per cell dry weight (CDW) were calculated using data from fluorescence measurement of whole bacterial cells as described in section II.6.9. All measurements were performed in at least three biological replicates. Biological variations were less than 10 %. **(b)** Soluble proteins of 5 x 10^8 cells taken from the indicated *B. megaterium* cultures were separated on 12 % SDS PAGE gels and visualized by Coomassie Brilliant Blue staining.

III.1.3 Combination of optimized promoter elements for protein production in *Bacillus megaterium*

A combination of the genetic elements enhancing target protein production was expected to cause even further increased protein yields. Therefore, new production plasmids were constructed combining two or more of the optimized DNA elements. First of all, the translation enhancing Shine Dalgarno sequence (*rbs*$^+$) was combined with every of the other optimized DNA sequences. Protein production experiments with Gfp as model protein were performed in *B. megaterium*. Cells harboring the plasmids pSSBm76 (-10$^+$/*rbs*$^+$) and pKMBm10 (*utr*$^+$/*rbs*$^+$) showed dramatically increased intracellular Gfp accumulations of 62.9 mg g_{CDW}^{-1} and 75.0 mg g_{CDW}^{-1}, respectively (Figure III.c). This increase is significantly higher than an addition of the protein amounts produced by the strains carrying the plasmids with only a single modification (-10$^+$: 20.4 mg g_{CDW}^{-1}, *utr*$^+$: 26.6 mg g_{CDW}^{-1}, *rbs*$^+$: 14.8 mg g_{CDW}^{-1}). The resulting increase in protein production was nearly a multiplication of the fold-changes observed for the single optimizations. Obviously, a higher amount or a more stable mRNA represents an improved target for the ribosome which is attracted by the perfect Shine Dalgarno sequence. Taken together, this combination yields in a highly synergistic protein synthesis process. In contrast, the combination of the optimized -35 promoter element (-35$^+$) and the adapted ribosome-binding site (*rbs*$^+$) did not show this strong effect. Cells harboring the plasmid pSSBm78 (-35$^+$/*rbs*$^+$) showed only a slightly increased Gfp content (80.6 mg g_{CDW}^{-1}) compared to cells employing pSSBm84 (-35$^+$) for protein production (76.5 mg g_{CDW}^{-1}). Analogously, the combination of both consensus promoter elements (-35$^+$/-10$^+$) with the optimized Shine-Dalgarno sequence (*rbs*$^+$) realized in the plasmid pSSBm81 did not yield the desired improvement in protein formation (16.7 mg g_{CDW}^{-1}). In fact, a strong decrease in Gfp production was detected, when compared with cells employing either pSSBm76 (-10$^+$/*rbs*$^+$) or pSSBm78 (-35$^+$/*rbs*$^+$) which encode a less conserved promoter. Ellinger *et al.* (1994) observed RNA polymerase stalling at promoters with very high similarity to the consensus during *in vivo* experiments in *E. coli*. Obviously, hampered by strong promoter affinity, the rate-limiting step was found to be the release of the sigma factor and thus the transition to an elongation complex (Ellinger *et al.*, 1994). Here, this stalling effect might have led to reduced transcription rates which consequently resulted in the formation of lower Gfp amounts. Further, under

non-inducing conditions detectable amounts of Gfp were found within *B. megaterium* cells harboring the full consensus promoter comprising plasmid pSSBm81 indicating a loss of promoter controllability (data not shown). Due to the highly increased affinity of the RNA polymerase complex to the consensus promoter (-35$^+$/-10$^+$), the XylR repressor binding might have lost its competitiveness. Since the binding sequence for the XylR repressor is located just a few base pairs downstream of P_{xylA}, it seems to be probable, that the major repressing effect of XylR is caused by competitive binding to the promoter region as it was shown for the *lac* repressor of *E. coli* by Schlax *et al.* (1995). *B. megaterium* cells producing Gfp mediated by a plasmid containing a combination of -35$^+$, *utr*$^+$ and *rbs*$^+$ showed a similar behaviour as cells harboring pSSBm81 (-35$^+$/-10$^+$/*rbs*$^+$) (data not shown). This might indicate that DNA sequence changes introduced to *utr*$^+$ do not exclusively influence the mRNA's half-life, but also significantly increase the RNA polymerase binding-affinity to its promoter.

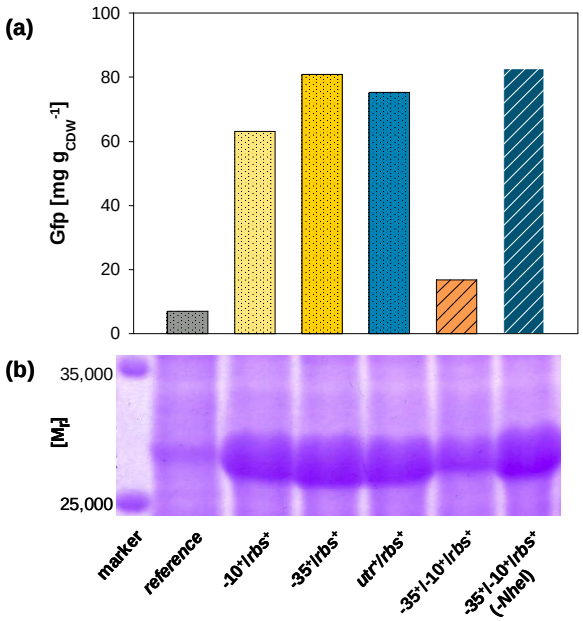

Figure III.c: Combinatory effect of optimized genetic elements on Gfp production.
B. megaterium MS941 cells were transformed with one of the indicated optimized expression plasmids harboring *gfp* under the control of the xylose-inducible promoter. The nature of combined genetically optimized elements (-35$^+$ region, -10$^+$ region, DNA sequence encoding the modified 5' *utr*$^+$ and the optimized ribosome binding site (*rbs*$^+$)) and the lack of the *Nhe*I restriction site are outlined. Cells carrying the non-optimized parental expression vector p3STOP1624-*gfp* were used as reference. Cultivations were performed in LB medium supplemented with appropriate antibiotics at 37 °C. Heterologous production of Gfp was induced with 0.5 % (w/v) xylose at OD_{578nm} = 0.4. Samples were taken 7.5 h after xylose addition. **(a)** Gfp amounts per cell dry weight (CDW) were calculated using data from fluorescence measurement of whole bacterial cells as described in section II.6.9. All measurements were performed in at least three biological replicates. Biological variations were less than 10 %. **(b)** Soluble proteins of 5 x 10^8 cells taken from the indicated *B. megaterium* cultures were separated on 12 % SDS PAGE gels and visualized by Coomassie Brilliant Blue staining.

Based on these results, an optimized expression plasmid was designed. Starting with the plasmid containing the combination of genetic elements (-35$^+$/*rbs*$^+$) facilitating the highest Gfp production (80.6 mg g_{CDW}^{-1}), the *Nhe*I site was changed back into its native DNA sequence since introduction of this unique restriction site revealed a negative impact on protein production in *B. megaterium*. Although, the introduction of the *Nhe*I restriction site into the non-optimized basic vector p3STOP1623 caused a significantly reduced Gfp production almost no differences in Gfp production were measured when comparing cells carrying the optimized plasmids with or without *Nhe*I

restriction site (pSSBm78: 80.6 mg g_{CDW}^{-1}, pSSBm85: 82.5 mg g_{CDW}^{-1}). These observations led to the speculation about general limitations of the protein formation process. On one site, this might be due to specific characteristics of the target gene. For example, the augmented use of certain rare codons in *B. megaterium* might lead to limitations of the corresponding tRNAs and could therefore hamper protein production. Another reason might be the limited supply of specific amino acids. This theory is supported by observations made during fed-batch cultivations. Here, on one hand high volumetric Gfp yields (1.25 g L^{-1}) were achieved, but on the other hand only lower Gfp amounts per cell dry weight (36.8 mg g_{CDW}^{-1}) compared to shaking-flask cultures (82.5 mg g_{CDW}^{-1}) were reached (Stammen *et al.*, 2010a). An explanation might be the use of defined synthetic medium in the fed-batch cultivations and complex medium employed in shaking flasks. Nevertheless, the combined optimized features facilitated the production of 82.5 mg g_{CDW}^{-1} Gfp in *B. megaterium* shaking-flask cultivations, which is even more than described for fed-batch cultivations of *E. coli* (73.7 mg g_{CDW}^{-1}) (Durrschmid *et al.*, 2008).

Thus, a genetic basis for competitive protein production using the alternative bacterial host *B. megaterium* was established. New high performance (hp) plasmids using the optimized elements -35^{+}/*rbs*$^{+}$ and the deletion of *Nhe*I site were created. For this purpose, combinations with coding sequences for N- or C-terminally localized His_6-tags for affinity chromatographical protein purifications were generated. An encoded TEV (tobacco etch virus) protease recognition site enables the removal of the N-terminally located His_6-tag. The novel hp-vectors p3STOP1623hp, pC-HIS1623hp and pN-HIS-TEV1623hp were constructed (Figure III.d) based on the established parental plasmids pSTOP1622, pC-HIS1622 and pN-HIS-TEV1622 (Biedendieck *et al.*, 2007c). Due to intensive investigation and documentation of the performance of these parental plasmids (Biedendieck *et al.*, 2007c) we refrained from somewhat redundant testing of the novel, fusion tag encoding hp-plasmids. The only difference to the here tested hp-vector equivalent for the intracellular Gfp production lies in the N- and C-terminally fused His_6-tags.

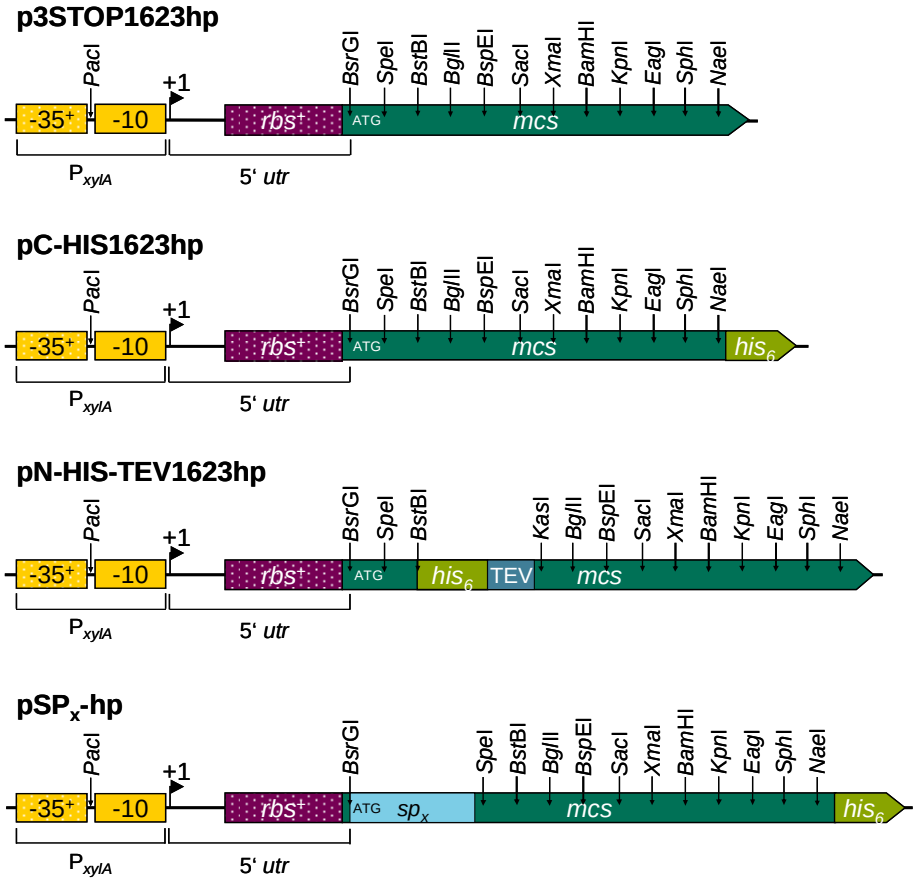

Figure III.d: Series of high performance (hp)-expression plasmids for recombinant protein production in *Bacillus megaterium*.

Schematic image of the core elements of the new hp-expression vectors p3STOP1623hp, pC-HIS1623hp, pN-HIS-TEV1623hp and the signal peptide (SP) encoding plasmids pSP_x-hp. The optimized promoter P_{xylA} (orange) consists of the optimized -35$^+$ and the native -10 promoter region. Between these promoter elements a restriction site for *Pac*I was introduced. Further, all of these high performance (hp)-plasmids encode a 5' untranslated region (5' *utr*) enclosing a Shine-Dalgarno sequence adapted to *B. megaterium* (*rbs*$^+$, violet). For simple subcloning all plasmids contain a comparable multiple cloning site (*mcs*, dark green). The most commonly used single cutting restriction enzymes are indicated. The *Bsr*GI site upstream of the start codon (ATG) enables the production of proteins with native N-terminus. For purification of intracellularly produced proteins via affinity chromatography, His_6-tags (green) can be C-terminally and N-terminally fused to the target protein employing the plasmids pC-HIS1623hp and pN-HIS-TEV1623hp, respectively. Additionally, after protein purification, the N-terminally located His_6-tag of pC-HIS-TEV1623hp can be removed by TEV (tobacco etch virus) protease treatment (TEV-recognition-site, blue). The signal peptide encoding plasmids pSP_x-hp enable the production and secretion of recombinant proteins in *B. megaterium*. SPs of the *B. megaterium* proteins YngK, Vpr, NprM, YocH, Pac, LipA and the artificial SP (sp_x, light blue) are available to facilitate secretion. For protein purification via affinity chromatography C-terminal fusions to His_6-tags can be constructed.

III.1.4 Identification of novel signal peptides for recombinant protein export in *Bacillus megaterium*

The signal peptides (SPs) for protein export via the Sec-pathway of *B. megaterium* were computationally predicted using the "PrediSi" software (http://www.predisi.de/) (Hiller *et al.*, 2004). During this process, all open reading frames (*orf*s) of the *B. megaterium* genome database "MegaBac" (version 2) were analyzed for signal peptide-coding sequences. The *orf*s had previously been determined using the software "Glimmer 2" (http://www.cbcb.umd.edu/software/glimmer/) (Delcher *et al.*, 1999). Most of the predicted *B. megaterium* signal peptides showed high similarity with the consensus sequence for type I signal peptides described for *B. subtilis* by Tjalsma *et al.* (2000; 2004). They had a length close to the predicted average of 28 amino acid residues and showed the typical three domains. The three leader sequences originated from the proteins Vpr (serine protease), NprM (neutral protease) and YngK (hypothetical protein) showed a high signal peptide probability and therefore were chosen to be examined. These proteins were found to be efficiently secreted by *B. megaterium* before (Wang *et al.*, 2006). However, the original DNA sequence of the signal peptide from YngK (sp_{yngK}) showed a poor codon adaptation index (CAI) of 0.33 for *B. megaterium* when analyzed with "JCat" software (http://www.jcat.de/) (Grote *et al.*, 2005). This low CAI indicated the use of codons, which are rare in *B. megaterium*. For production of *T. fusca* hydrolase (Tfh) in *B. megaterium* it was shown, that the use of rare codons in a coding sequence may lead to a slowdown of the translational process or to complete abortion of translation (Yang *et al.*, 2007). Thus, additionally to the native sp_{yngK}, a codon-optimized version sp^{+}_{yngK} (CAI = 0.99) was synthesized *de novo* and tested. The signal peptide of YocH (hypothetical protein) was chosen from the list of predicted *B. megaterium* SPs due to its very high similarity to the consensus signal peptide structure for Sec-dependent protein secretion described by Tjalsma *et al.* (2000; 2004). So far, the protein YocH was not detected in any secretome analysis of *B. megaterium*. In addition to these native *B. megaterium* SPs, an artificial signal peptide (SP_{Asp}) was constructed. This *in silico* designed "ideal" signal peptide was deduced from the average amino acid sequence of all signal peptides currently known for Gram-positive bacteria. A hidden Markov model approach implemented in the program PrediSi (Hiller *et al.*, 2004) was used for the development of this peptide.

Based on a systematic database search, the amino acid residue for each position was chosen which occurs most frequently at that position. Two SPs already used in recombinant protein production and secretion in *B. megaterium* served as positive control in this work. The signal peptide of penicillin amidase (SP_{Pac}) (Panbangred *et al.*, 2000) and of the esterase LipA (SP_{LipA}) (Ruiz *et al.*, 2002) were already tested to facilitate the secretion of a heterologous levansucrase LevΔ773 in *B. megaterium* (Biedendieck *et al.*, 2007a). The SP_{LipA} was also successfully employed in *B. megaterium* to mediate the efficient export of a hydrolase from *T. fusca* (Yang *et al.*, 2007).

Table III.a: New *Bacillus megaterium* signal peptides evaluated in this work.

The nucleic acid (small letters) and amino acid sequences (capitals) as well as the codon adaptation indices (CAI) of the novel signal peptides used in this work are shown in this table. The leader peptide sequences are separated from the N-terminus of the corresponding mature protein sequences by an apostrophe. The bases different between sp_{yngK} and its codon adapted pendant (sp^{+}_{yngK}) are underlined.

Name	CAI	Nucleic and amino acid sequence
sp_{vpr} SP_{Vpr}	0.65	atgaaaaaagtatcaattcggtctgtattaagtacagtagcattttcgg ttgctttatcttctttgcaatgggagcatccgcc'aac MKKVSIRSVLSTVAFSVALSSFAMGASA'N..
sp_{nprM} SP_{NprM}	0.62	atgaaaaagaaaaaacaggctttaaaggtattattatcagttggtatcc tttcttcatcatttgcttttgcacatacgagcagtgcc'gcg MKKKKQALKVLLSVGILSSSFAFAHTSSA'A..
sp_{yocH} SP_{YocH}	0.46	atgaaaaagacaatgattacgttcagtttagttctaatgtcactgttcg gagtggcaagtggtgcatcagct'gca MKKTMITFSLVLMSLFGVASGASA'A..
sp_{yngK} SP_{YngK}	0.33	atgtatattaaaaaatgtataggaagcatcttgtttttactgctgtttt gttcgagcgcgttacctgctaaagca'gat MYIKKCIGSILFLLLFCSSALPAKA'D..
sp^{+}_{yngK} SP^{+}_{YngK}	0.99	atgtatattaaaaaatgtattggttctattttattcttattattattct gttcttctgctttaccagctaaagct'gat MYIKKCIGSILFLLLFCSSALPAKA'D..
sp_{asp} SP_{Asp}	0.81	atgcgtaaaagaaaaaaagcattagctgtagcattactttagcagcat taggtgcattacttccaacaacagcatctgct'gat MRKRKKALAVALLLAALGALLPTTASA'D..

III.1.5 Recombinant protein export driven by novel signal peptides from *Bacillus megaterium*

All new signal peptides coding regions (Table III.a) as well as the sp_{lipA} and sp_{pac} were introduced into both the plasmid containing the native xylose-inducible expression system and the plasmid harboring the improved hp-expression system. All plasmids were equipped with a codon adapted variant of the *tfh* gene fused to the coding region of a C-terminally located His_6-tag (Yang *et al.*, 2007). For Tfh a simple and well reproducible enzyme activity assay is available, which makes the exoenzyme a viable model for heterologous protein production and secretion in *B. megaterium* (Yang *et al.*, 2007).

In production experiments it was shown that all employed SPs facilitate secretion of active heterologous Tfh into the growth medium (Figure III.e). Culture supernatant of *B. megaterium* cells transformed with a plasmid without *tfh* gene (pMM1525) was used as negative control and did not show any hydrolase activity at all. Tfh protein amounts quantified via SDS-PAGE gels correlated well with the detected volumetric hydrolase activities indicating that all secreted enzyme was in its active conformation (compare Figure III.e, panel (a) with panel (b)).

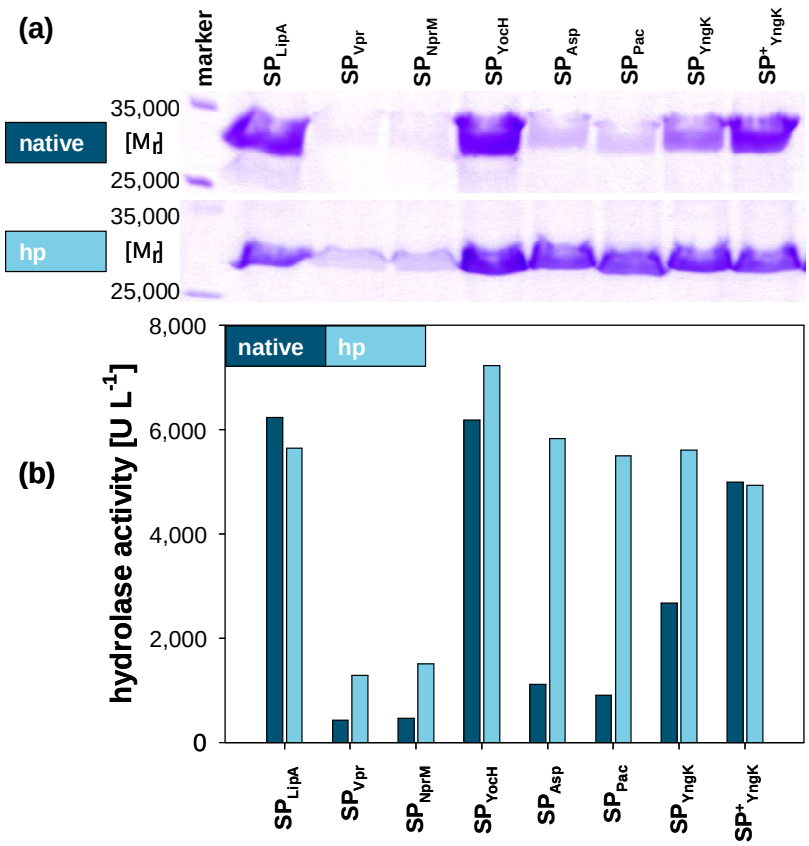

Figure III.e: Secretory production of Tfh facilitated by new signal peptides.

B. megaterium MS941 cells were transformed with either the native or the hp-plasmids encoding the indicated SPs and Tfh-His_6. "native" stands for the P_{xylA} driven expression system (Malten *et al.*, 2006) and "hp" stands for the high performance plasmid system as described in section III.1.3 of this work. Cultivations were performed in LB medium supplemented with the appropriate antibiotics at 37 °C. Heterologous production of Gfp was induced with 0.5 % (w/v) xylose at OD_{578nm} = 0.4. Samples were taken 9 h after xylose addition. **(a)** Proteins from 1.5 mL cell-free supernatant from *B. megaterium* cultures carrying plamids mediating Tfh export by the indicated leader peptides were precipitated with ammonium sulfate. Subsequently, proteins were separated on 12 % SDS PAGE gels and visualized by Coomassie Brilliant Blue staining. **(b)** Hydrolase activities were determined photometrically at 30 °C by the *p*-nitrophenylpalmitate (*p*NPP) method as described in section II.6.10. All measurements were performed in at least three biological replicates. Biological variations were less than 10 %.

Generally, *B. megaterium* cells employing the high performance (hp)-production system (pSSBm94 - pSSBm101) showed increased amounts of Tfh secreted into the growth medium compared to cells employing the corresponding basic expression systems (pSSBm28 - pSSBm31, pSSBm33 - pSSBm35 and pYYBm9). Used in combination with the basic expression plasmids, the four signal peptides SP_{Vpr},

SP_{NprM}, SP_{Pac} and SP_{Asp} facilitated secretion of low Tfh amounts (SP_{Vpr} = 430 U L^{-1}, SP_{NprM} = 470 U L^{-1}, SP_{Pac} = 910 U L^{-1} and SP_{Asp} = 1,100 U L^{-1}). The signal peptide SP_{YngK} mediated export of medium quantities of the model enzyme (SP_{YngK} = 2,700 U L^{-1}) and the three leader sequences SP^{+}_{YngK}, SP_{YocH} and SP_{LipA} facilitated the transport of high Tfh amounts into the growth medium (SP^{+}_{YngK} = 5,000 U L^{-1}, SP_{YocH} = 6,200 U L^{-1} and SP_{LipA} = 6,200 U L^{-1}) (Figure III.e). Comparison of the systems employing a native sp_{yngK} or a codon adapted signal peptide sp^{+}_{yngK} showed, that significantly more Tfh is produced and secreted by cells harboring an adjusted signal peptide (SP_{YngK} = 2,700 U L^{-1}, SP^{+}_{YngK} = 5,000 U L^{-1}). This underlines the importance of the codon usage for the translation, which seemed to be the rate limiting step within the production and secretion process when using the basic expression system. No significant advantage in the protein production and translocation could be detected by using the codon adapted coding region of the signal peptide SP_{YngK} as can be determined in the data using the hp-expression plasmids (SP_{YngK} = 5,600 U L^{-1}, SP^{+}_{YngK} = 4,900 U L^{-1}). When employing either SP_{Asp} or SP_{Pac} in combination with the hp-expression system in *B. megaterium* secreted Tfh amounts in the culture supernatants were dramatically increased compared to the basal system. For SP_{Asp} the increase was > 5-fold (5,800 U L^{-1}) while for SP_{Pac} even a 6-fold increase (5,500 U L^{-1}) was detected. Still a significant increase of exported Tfh was detected in the culture broth of cells which carried the hp-expression system in combination with the signal peptides SP_{Vpr} (1,300 U L^{-1}), SP_{NprM} (1,500 U L^{-1}) and SP_{YngK} (5,600 U L^{-1}). The three leader peptides, which performed well in combination with the basic expression system facilitated secretion of comparable Tfh amounts when used in the hp-system (SP^{+}_{YngK} = 4,900 U L^{-1}, SP_{YocH} = 7,200 U L^{-1} and SP_{LipA} = 5,600 U L^{-1}).

Dispite the use of the new hp-plamids and the testing of several novel signal peptides only a 16 % higher secretion of Tfh was detected (SP_{YocH} in hp-plasmid: 7,200 U L^{-1} equals 7.7 mg L^{-1}) compared to the one mediated by SP_{LipA} used in combination with the non-optimized expression system (6,200 U L^{-1} or 6.6 mg L^{-1}). These results clearly show limitations of the Sec-dependent secretion process, since the intracellular protein production was drastically increased up to 18-fold (4.6 vs. 82.5 mg g_{CDW}^{-1}) as shown for Gfp production when employing the hp-production system. Further, it was shown that the computational approach to create an artificial signal peptide is possible in general. Application of a SP_{Asp} containing expression

plasmid in protein secretion experiments proved the sufficient biological functionality of the artificial leader peptide. Since no prediction method for the right combination of SP and target protein is currently applicable with success (Brockmeier *et al.*, 2006), screening for the correct SP-protein combination is still necessary. Nevertheless, in combination with the newly developed hp-expression system six (SP_{YocH}, SP_{YngK}, SP^{+}_{YngK}, SP_{Asp}, SP_{Pac}, SP_{LipA}) out of eight tested SPs (75 %) facilitated high yield secretion of Tfh, while only three SPs (SP_{YocH}, SP^{+}_{YngK}, SP_{LipA}) (38 %) achieved an equivalent performance when employed in the non-optimized production system. Thus, when using the new hp-plasmids, less affort is necessary to screen for an efficiently working SP-protein combination.

For future protein production applications, the established combination of *mcs* and coding sequence for a C-terminally located His_6-tag of pHIS1525 (Malten *et al.*, 2006) was inserted into the high performance production plasmids comprising one of each of the presented SPs (pSSBm94 - pSSBm101) (Figure III.d). A desired target protein can be cloned into these plasmids generating His_6-tagged proteins, which are produced and secreted into the medium by *B. megaterium* and which can subsequently be purified from the supernatant by metal-affinity chromatography as demonstrated by Malten *et al.* (2006). Many of the hp-expression plasmids for both intracellular and extracellular protein production described in chapter III.1 of this work are commercially available by MoBiTec GmbH, Göttingen, Germany.

III.2 New P_{xylA}-based production plasmid for *Bacillus megaterium* designed for easy cloning

The productivity of the most commonly employed gene expression system based on the xylose-inducible promoter P_{xylA} of the *B. megaterium xyl*-operon was improved in section III.1 of this work. Now, a broad plasmid toolbox enables the high yield production of both recombinant intra- and extracellular proteins. However, sometimes cloning of target genes under control of P_{xylA} in *E. coli* causes problems. It is supposed, that major complications within the cloning host *E. coli* are based on the non-controllable transcription initiation at the, in *B. megaterium* tightly controllable P_{xylA} promoter. This leads to a constant basal production of the target protein during establishment of newly constructed plasmids in *E. coli*. Production of some proteins has negative or even cytotoxic effect on bacterial growth and via negative selective pressure this leads to plasmids encoding non- or less functional proteins. Thus, even though the final production in *B. megaterium* might be feasible, projects already fail due to difficulties during vector amplification in *E. coli*.

To circumvent these problems diverse fusion promoters were designed to make the P_{xylA} promoter controllable during cloning in *E. coli* (data not shown). Since all these attempts failed, another strategy was followed. In this section of the work, a plasmid hybrid was constructed containing the approved optimized *Bacillus* elements (Stammen *et al.*, 2010a) in combination with *E. coli* elements of the pBAC / *oriV* (Wild *et al.*, 2002). These enable the maintenance of a plasmid in single-copy state intended for the use during the cloning procedure in *E. coli*. By reduction of the plasmid copy number the level of basal expression initiated at P_{xylA} and thus the amount of protein with potentially toxic effect on *E. coli* can be limited to a niveau enabling the growth of positive transformants. Short time before plasmid preparation, the inducible origin of replication (*oriV*) can be activated by an up-variant of TrfA. Thus, by expression of the *trfA* gene, the plasmid copy number and consequently the plasmid DNA yields can be increased.

III.2.1 Construction of the Ecoco system

The basis for the *Bacillus* parts of the new plasmid system named “Ecoco” (*Escherichia coli* copy control) was the high performance expression plasmid p3STOP1623hp (section III.1). Its elements for the xylose-inducible protein production and replication in *Bacillus* as well as the tetracycline resistance cassette were used. For maintenance in single-copy state in *E. coli* the *oriS* origin of replication, the *repE* gene and *parABC* partition determinants were amplified via PCR from a pBAC / *oriV* template vector and inserted in the Ecoco-plasmid. Further, *oriV* for maintenance in multi-copy state and the β-lactamase gene for resistance-based selection in *E. coli* of pBAC / *oriV* were used. At the end of this cloning procedure, a plasmid named pBmEcoco was generated (Figure III.f). It combined the highly functional elements for controlled protein production in *B. megaterium* with the elements for single-copy number maintenance. Further, it harbors the *oriV*, which is responsible for the amplification of pBmEcoco to multi-copy state in *E. coli*. To reduce the size of pBmEcoco the gene encoding the TrfA up-mutant necessary to induce the replication at *oriV* was placed on a second, auxiliary plasmid. This additional vector is just necessary during the cloning and amplification phase of pBmEcoco in *E. coli*. On this pBAD33 derivative, the *trfA* gene was inserted under control of the pBAD promoter. *E. coli* cells harboring this new plasmid named pEc-*trfA* produce TrfA when arabinose is present in the medium and thereby induce the multi-copy state of the pBmEcoco plasmid (Figure III.f).

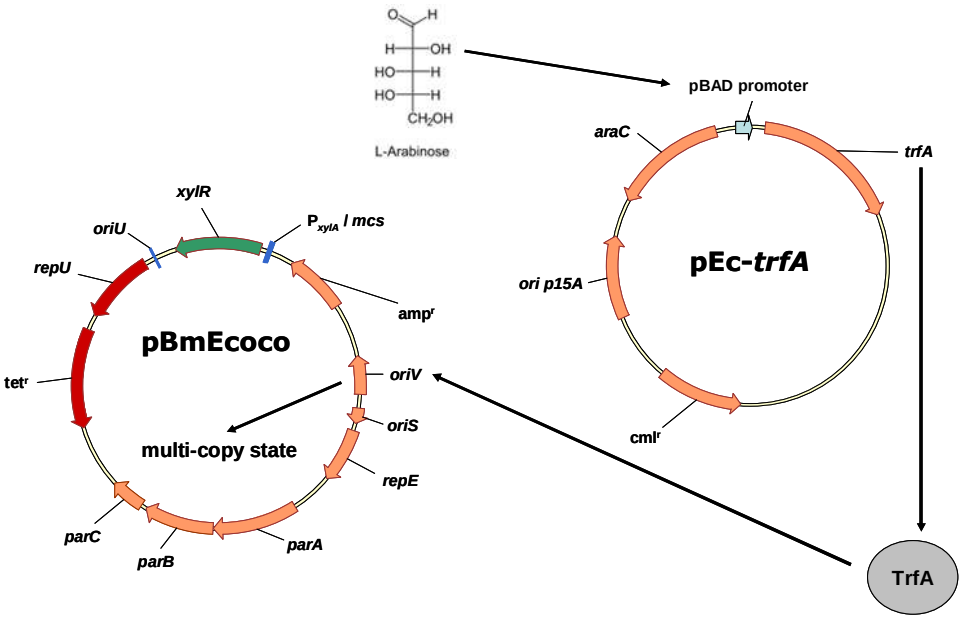

Figure III.f: Funtionality of the Ecoco system in *Escherichia coli*.
Both plasmids (pBmEcoco and pEc-*trfA*) of the Ecoco system are present in *E. coli* cells during cloning and amplification of new pBmEcoco constructs. Elements with function in *E. coli* are colored in orange. The auxillary plasmid pEc-*trfA* carries the *trfA* gene under control of the arabinose inducible promoter pBAD (light blue). Further, AraC, the constitutively produced repressor of pBAD, is encoded on pEc-*trfA* in reverse orientation to the pBAD promoter. Addition of arabinose leads to conformational changes in AraC and transcription of *trfA*. Plasmid maintenance of pEc-*trfA* is controlled by the *p15A* origin of replication (*ori*). A chloramphenicol resistence (cml^r) enables selection on plasmid carrying bacteria. Next to the elements functional in *Bacillus* (P_{xylA} / *mcs*, *xylR*, *oriU*, *repU* und tet^r), pBmEcoco contains genetic information for maintenance in single-copy state (*parABC*, *repE* and *oriS*) in *E. coli*. An ampicillin resistance gene (amp^r) enables selection of plasmid containing *E. coli* cells. Induction of *oriV* by the up-mutant protein of TrfA leads to the upregulation of the pBmEcoco copy number to multi-copy state.

III.2.2 Verification of the "on command" amplification of plasmid copy number in *Escherichia coli*

To verify the functionality of the novel Ecoco system, the *gfp* gene was inserted into the multiple cloning site (*mcs*) of pBmEcoco under the control of P_{xylA}. Since P_{xylA} shows leaky basal expression of a gene under its control in *E. coli*, bacterial cells harboring the plasmid in several copies appear light green under blue light, when using a yellow filter. This fluorescence phenomenon originated from Gfp accumulation inside the bacterial cells was already observed during plasmid

preparations in section III.1. Hence, the fluorescence intensity correlated directly with the plasmid copy number. *E. coli* DH10B cells harboring the auxiliary-plasmid pEc-*trfA* and pBmEcoco-*gfp* were cultivated in LB medium supplemented with the appropriate antibiotics (chloramphenicol 17 µg mL^{-1} and carbenicillin 50 µg mL^{-1}) with or without additional arabinose (0.02 % (w/v)). Eight hours after inoculation, when bacteria started to enter stationary growth phase, Gfp fluorescence was measured. *E. coli* cells harboring the appropriate *gfp* encoding plasmid of the hp-plasmid series (pSSBm85) cultivated in LB medium with carbenicillin supplementation were used as reference. Bacterial cells harboring the Ecoco system and cultivated in the presence of arabinose showed comparable Gfp amounts like *E. coli* cells carrying the reference plasmid pSSBm85 (Figure III.g panel (b)). This was already assumed by visually comparison of the bacterial cultures under blue light (Figure III.g panel (a)) and this result was verified by the fluorescence measurement. Furthermore, fluorescence could neither visually nor analytically be determined in cells carrying the Ecoco system, which were grown without supplementary arabinose.

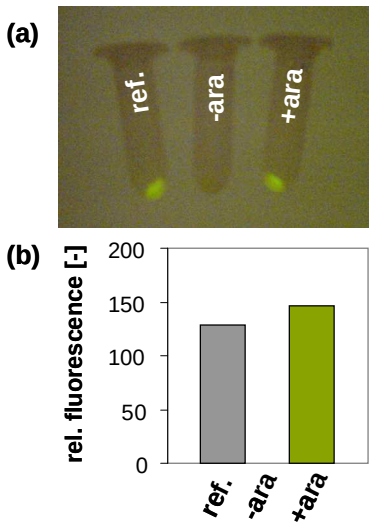

Figure III.g: Relative quantification of the plasmid copy number by Gfp fluorescence analysis.
E. coli DH10B cells were transformed either with the *gfp* carrying Ecoco system (pBmEcoco-*gfp* and pEc-*trfA*) or the *gfp* harboring optimized xylose-inducible system (ref.) described in section III.1 (pSSBm85). Bacteria were cultivated aerobically in LB medium supplemented with the appropriate antibiotics at 37 ℃. Either 0.02 % arabinose was a dded (+ara) or was not added (-ara) to the media of *E. coli* cells carrying the plasmids of the Ecoco system. Samples were taken approximately 8 hours after inoculation. **(a)** 3 x 10^9 cells were sedimented in reaction tubes by centrifugation and Gfp fluorescence was visualized on a blue screen using a yellow filter. **(b)** Relative Gfp fluorescence was measured in whole bacterial cells as described in II.6.9.

Additonally, plasmid DNA was prepared from these Ecoco system carrying *E. coli* cells. Appropriate amounts were digested with the restriction endonuclease *Sph*I. This enzyme is a single cutter of both the pBmEcoco-*gfp* and the auxiliary vector pEc-*trfA*. After agarose gel electrophoreses of the linearized plasmids, a significantly higher amount of pBmEcoco-*gfp* was found for the sample originated from the cultivation with arabinose supplementation (Figure III.h). Moreover, the band originated from pEc-*trfA* represents a smaller amount of the auxiliary plasmid. Due to differences in plasmid preparations, this implied an even higher difference in the pBmEcoco-*gfp* copy number.

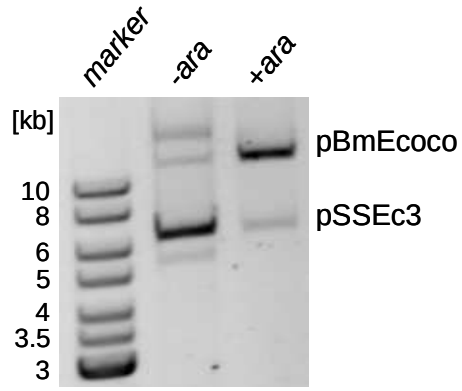

Figure III.h: Comparison of the cellular plasmid DNA content.
E. coli DH10B cells transformed with the *gfp* carrying Ecoco system (pBmEcoco-*gfp* and pEc-*trfA*) were cultivated aerobically in LB medium supplemented with the appropriate antibiotics at 37 °C. The medium was either supplemented with 0.02 % arabinose (+ara) or not (-ara). Plasmid preparations were done approximately 8 hours after inoculation. The GeneRuler DNA ladder from Fermentas was used as size standard (marker). Bands corresponding to the linearized plasmids pBmEcoco and pEc-*trfA* are indicated.

III.2.3 Performance of the novel expression plasmid in *Bacillus megaterium*

III.2.3.1 Intracellular recombinant production of green fluorescent protein

After “on command” amplification of plasmid copy number was proven to be functional in *E. coli*, the pBmEcoco-plasmid was tested for its functionality in *B. megaterium*. For this purpose, the Gfp encoding plasmid pBmEcoco-*gfp*

introduced in section III.2.2 was used for protein production experiments in *B. megaterium* MS941. Cells of *B. megaterium* harboring the plasmid p3STOP1624-*gfp* (*gfp* expressed by the native xylose-inducible protein production system) or pSSBm85 (*gfp* expressed by the optimized hp-protein production system) were used as reference for Gfp production.

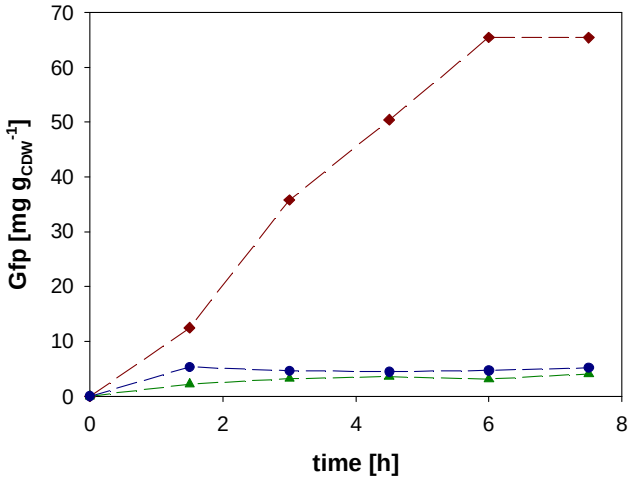

Figure III.i: Gfp production in *Bacillus megaterium* using the Ecoco system

B. megaterium MS941 cells were transformed with the plasmid pBmEcoco-*gfp* (dark blue), the corresponding hp-plasmid pSSBm85 (P_{xylA}-(-35^{+}+*rbs*$^{+}$)-*gfp*) (dark red) or the basic expression plasmid encoding *gfp* (p3STOP1624-*gfp*; P_{xylA}-*gfp*) (green). These strains were cultivated aerobically in LB medium supplemented with the appropriate antibiotics at 37 °C. Heterologous production of Gfp was induced with 0.5 % (w/v) xylose at OD_{578nm} = 0.4. Samples were taken every 1.5 h after xylose addition. Gfp amounts per cell dry weight (CDW) were calculated using data from fluorescence measurement of whole bacterial cells as described in section II.6.9.

Surprisingly, bacteria of cultures carrying the plasmid pBmEcoco-gfp showed a much lower Gfp content (5.2 mg g_{CDW}^{-1}) than cells employing pSSBm85 for protein production (Figure III.i). This was astonishing, since the elements for replication, antibiotic resistance and protein production for *B. megaterium* of both plasmids are completely identical. Thus, we expected a comparable plasmid copy number and a comparable Gfp production. Since fluorescence microscopic analysis unveiled a significant increase of cells not producing Gfp, it was supposed that the pBmEcoco vector might be instable in *B. megaterium* and got lost during the cultivation. However, plating and counting of resistant and hence plasmid-carrying cells was not feasible due to a broad colony size distribution. While small colonies were barely

visible, large ones already overgrew others. However, even though, the protein production is reduced, it still exceeds the productivity of the native P_{xylA}-based expression system (Figure III.i).

III.2.3.2 Production and secretion of the hydrolase of *Thermobifida fusca*

Additionally, the Ecoco system was tested for the production and secretion of target proteins. As in section III.1.5 the hydrolase of *T. fusca* (Tfh) was used as model protein. The gene 5'-fused to the coding sequence of the SP_{YocH} was inserted into the *mcs* of the pBmEcoco plasmid. *B. megaterium* MS941 transformed with this plasmid was employed in production experiments. *B. megaterium* harboring pSSBm97, the appropriate pendant of the hp-plasmid series encoding Tfh was used as reference.

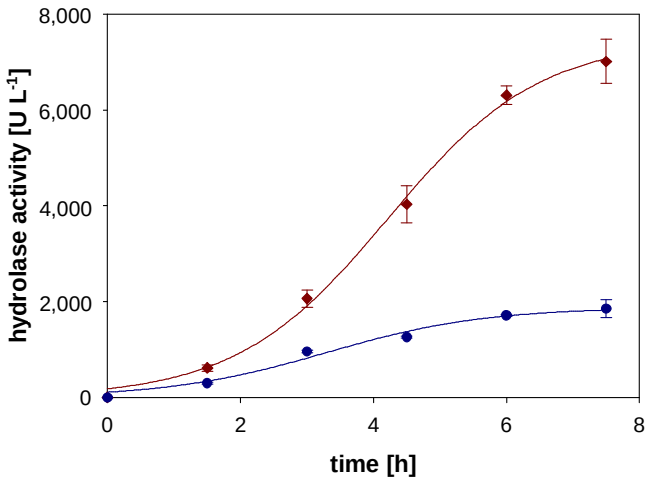

Figure III.j: Production and secretion of the Tfh applying the Ecoco system in *Bacillus megaterium*

Cells of the *B. megaterium* strain MS941 were transformed with the pBmEcoco plasmid encoding a SP_{YocH}-Tfh-His_6 fusion protein (dark blue). As reference (dark red) the corresponding hp-plamids (P_{xylA}-(-35$^+$+*rbs*$^+$)-sp_{yocH}-*tfh*-his_6) was employed in *B. megaterium* MS941. These strains were cultivated aerobically in LB medium supplemented with the appropriate antibiotics at 37 °C. Heterologous production and secretion of Tfh was induced with 0.5 % (w/v) xylose at OD_{578nm} = 0.4. Samples were taken every 1.5 h after xylose addition. Hydrolase activities were determined photometrically at 30 °C by the *p*-nitrophenylpalmitate (*p*NPP) method as described in section II.6.10.. All measurements were performed in three biological replicates. Standard deviations are indicated as error bars.

Results of the production and secretion experiments using the model protein Tfh revealed, that *B. megaterium* with the Ecoco system produced less Tfh when compared to bacteria employing the hp-system (Figure III.j). Finally, after 7.5 h only 26 % of the Tfh amount generated and secreted by the hp-system (7,000 U L^{-1}) was produced and transported into the medium by the Ecoco system (1,800 U L^{-1}). Since a reduced productivity of every individual cell would not lead to this significant reduction in secreted Tfh amounts, these results strengthen the hypothesis of plasmid instability and subsequent plasmid loss as already constructed in section III.2.3.1. The outcome of decreased productivity on single cell level was shown when the secretory production of Tfh by the native P_{xylA}-based expression system was compared to the production and secretion by the optimized hp-expression system (section III.1.5). The Tfh amounts secreted by *B. megaterium* cells showed no significant difference between the two systems even though the intracellular protein formation of the hp-system was much higher as it could be shown for the formation of Gfp (section III.1.3). However, data revealed that the new Ecoco system is functional for production and secretion of target proteins with *B. megaterium*.

III.3 K1E Phage RNA polymerase: a new tool for *in vitro* RNA synthesis and protein production *in vivo*

Only a limited number of phage RNA polymerases are currently commercially available for *in vitro* and *in vivo* purposes. They differ in their modes of promoter recognition (Jorgensen *et al.*, 1991), transcription initiation (Chakraborty *et al.*, 1977) and enzymatic activity within the host cell, which may not always be compatible with the envisaged RNA or protein production process (Stump & Hall, 1993). Moreover, further developments of existing systems are sometimes hampered by intellectual property constraints (*e.g.* patents (McAllister, 1991; Studier *et al.*, 1999)). Consequently, a broader spectrum of different RNAPs for *in vitro* and *in vivo* applications is of interest.

In this part of the work, the yet non-characterized DNA-dependent RNAP of the enterobacteriophage K1E was made accessible to high yield *in vitro* transcription and protein production in *B. megaterium*. Subsequently, new K1E RNAP-based production systems were compared to both the existing T7 gene expression system for *B. megaterium* (Gamer *et al.*, 2009) and the also newly developed protein production system for *B. megaterium* employing the RNAP of *S. typhimurium* phage SP6 (Stammen *et al.*, 2010b).

III.3.1 Recombinant production and purification of K1E RNA polymerase

Bacteriophage K1E derived RNAP was recombinantly produced in *E. coli* BL21 (DE3) cells. The His_6-tagged fusion protein was purified by metal-affinity chromatography and concentrated up to 12 mg mL^{-1}. Over 95 % of the obtained protein was the desired RNAP (M_r = 100,000) as shown by SDS-PAGE gel analysis (Figure III.k). A volumetric enzyme amount of 22 units μL^{-1} and a RNAP content of 95 % led to the determination of the specific RNAP activity of 4,600 units mg^{-1}.

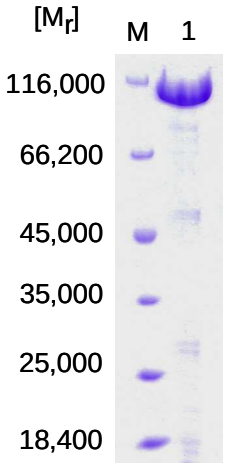

Figure III.k: Evaluation of K1E RNA polymerase purity via SDS-PAGE gel analysis.
K1E RNA polymerase was heterologously produced in *E. coli* BL21 (DE3) cells. Bacterial cells harboring the plasmid pET-32a(+)-K1EP were cultivated aerobically at 37 °C. Production of heterologous RNAP was induced at OD_{578nm} = 0.5 by addition of 100 µM IPTG to the growth medium. Four and a half hours after induction of heterologous gene expression, cells were harvested by centrifugation and disrupted. After removal of cell debris by ultra-centrifugation the supernatant was subjected to metal-affinity chromatographie. RNAP activity was assayed as described in section II.6.11. Fifty units of purified K1E RNA polymerase were applied to a SDS-PAGE gel (lane 1). Purity of the enzyme fraction was visually estimated. The unstained protein molecular weight marker from Fermentas was used as size standard (lane M).

E. coli BL21(DE3) cells were transformed with the plasmid pET-32a(+)-K1EP and employed for production of the His_6-tagged K1E phage originated RNAP. Baffled shaking flasks with LB medium were inoculated with 5 x 10^7 cells of a starter-culture and subsequently cultivated at 37 °C with vigorous shaking. At OD_{578nm} = 0.5, Isopropyl-β-D-thiogalactopyranosid (ITPG) was added to the growth medium to a final concentration of 100 µM. Approximately 4.5 h after induction of heterologous gene expression cells were sedimented at 4 °C and the target protein was purified after cell disruption (see section II.6.6).

III.3.2 Identification of K1E RNA polymerase promoters

The genome of *E. coli* phage K1E was sequenced by Stummeyer *et al.* (2006), but no promoters for the K1E RNAP were annotated. Thus, it was necessary to screen the phage genome for possible promoter sequences. For this purpose, promoter

sequence logos (Schneider & Stephens, 1990) of RNAPs from SP6-like phages created by Chen and Schneider (2005) were used as template to search the K1E genome for putative phage RNAP promoters employing the software "Virtual Footprint" (Münch *et al.*, 2005). Surprisingly, using the SP6 promoter logo as template no matching DNA sequences in the K1E genome were found. The use of the promoter logo of *E. coli* bacteriophage K1-5 RNAP (Chen & Schneider, 2005) for a K1E genome wide promoter search led to 13 sequence hits (Table III.b). This was not unexpected, since the K1E genome showed 92 % identity to its K1-5 counterpart. The computationally predicted putative K1E promoters were manually re-evaluated with respect to their precise upstream location to correspondingly annotated genes. Twelve of the 13 predicted potential promoter sequences matched the criterion of appropriate promoter location. These sites were subsequently used to calculate a promoter logo for the RNAP of bacteriophage K1E (Figure III.l) employing the web application "Weblogo 3.0" (Crooks *et al.*, 2004).

Table III.b: Predicted promoter sequences within the genome of bacteriophage K1E.
The putative promoter sequences found in the genome of *E. coli* phage K1E were predicted using the software "Virtual Footprint" (Münch *et al.*, 2005) and the promoter logo of bacteriophage K1-5 (Chen & Schneider, 2005) as template.

Sequence (+1 site bold)	**Location** (bp)	**Upstream of gene**
actatttagctgacactata**a**gagaa	2,776 - 2,801	5
tttatttaccggacactata**g**gatag	6,821 - 6,846	-
cgtatttaccggacagtata**g**ataag	7,005 - 7,030	8
cggatttaccggacactata**g**aagag	9,339 - 9,364	9.5
atcatttcgccgacactata**g**aaggt	10,187 - 10,212	10.5
atgatttactggacactata**g**aagga	13,922 - 13,947	15
agtaattactggacactata**g**aacaa	14,691 - 14,716	16
ggtatttactggacactata**g**aagag	15,929 - 15,954	18
aagatttagttgacactata**g**aacaa	19,094 - 19,119	23
agtatttactggacactata**g**aaggg	24,428 - 24,453	31
aggcaatactggacactata**g**aagaa	37,179 - 37,204	40A
attaattaccggacactata**g**aagga	40,030 - 40,055	41
acgatttactggacactata**g**aagag	41,714 - 41,739	46

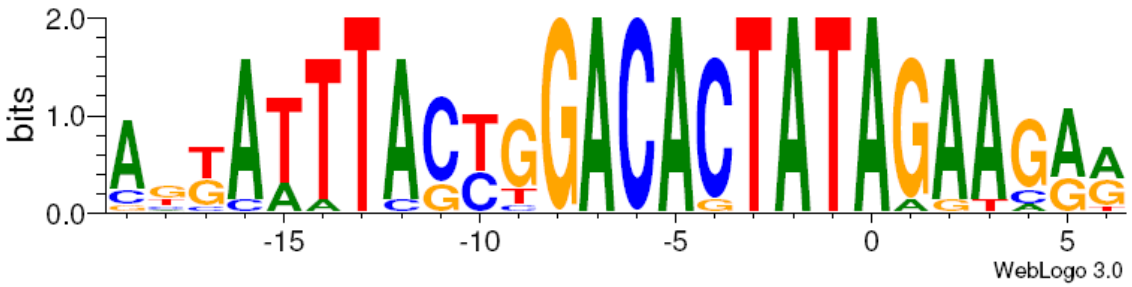

Figure III.l: Sequence logo of the K1E RNA polymerase promoter site.
A promoter logo for K1E RNAP was created with "weblogo 3.0" (Crooks *et al.*, 2004) using twelve putative promoter sites. No adjustment of the base composition was applied. The bases of the region -19 to +6 regarding the predicted transcription start site are shown.

Further, two of these predicted promoters were analyzed for their potential to initiate transcription of a target DNA sequence by K1E RNA polymerase *in vitro*. One promoter (K1E-1: atttactggacactata**g**aa; +1 base in bold) perfectly matched the core region (-17 - +3) of the predicted consensus (Figure III.l). The other (K1E-2: atttagctgacactata**a**ga; +1 base in bold) revealed a high homology to the SP6 RNAP promoter (atttaggtgacactata**g**aa; +1 base in bold). Further, it was characterized by an adenine instead of the conserved guanine residue at position +1. This guanine residue was found to be essential for significant transcription initiation at phage RNAP promoters (Milligan *et al.*, 1987; Stump & Hall, 1993). The three described promoters, K1E-1, K1E-2 and SP6 were placed individually upstream of the plasmid-localized model gene *gfp*. Subsequently, the *gfp*-mRNA was the target for *in vitro* transcription assays.

III.3.3 *In vitro* transcription using K1E RNA polymerase

Since RNAP of enterophage K1E showed an amino acid identity of 84 % to the RNAP of *Salmonella* phage SP6, it was concluded that its assay requirements might be similar to the ones of SP6 RNAP, which is commercially available for *in vitro* transcription applications. Thus, first activity assays were performed using similar reaction conditions as described by Melton *et al.* (1984) for *in vitro* transcription using SP6 RNAP. The *gfp*-RNA amounts were quantified with the RNA Pico 6000 Kit and the Agilent Bioanalyzer 2100. Significant amounts of *gfp*-RNA were synthesized

under the tested conditions by K1E RNAP when transcription was initiated at the K1E-1 promoter. The linear correlation of used K1E RNAP amounts and the detected RNA levels was demonstrated.

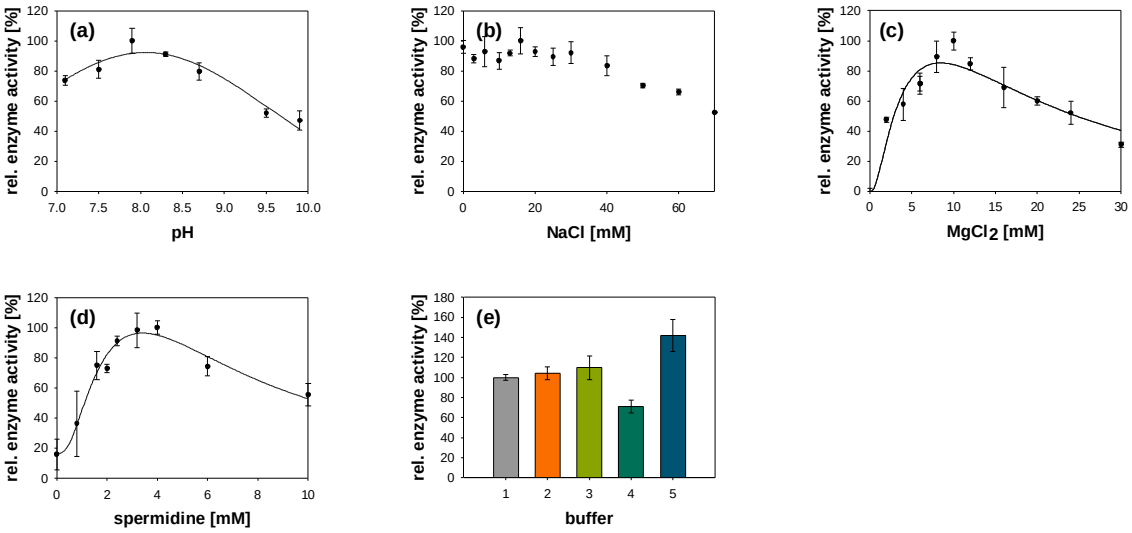

Figure III.m: Establishment of *in vitro* transcription conditions for the RNA polymerase of K1E bacteriophage.

In vitro transcriptions using K1E RNAP were performed as outlined in section II.6.11. Parameters modified (**(a)** pH-value, **(b)** NaCl, **(c)** $MgCl_2$ and **(d)** spermidine) are indicated on the x-axis of the diagrams. Maximum RNA yields corresponded to 100 % relative enzyme activity. All other values were related to that. In panel **(e)** RNA amounts produced within the non-optimized buffer (lane 1) were set to 100 %. The buffers of the various assays contained 6 mM $MgCl_2$ and 2 mM spermidine (lane 1), 10 mM $MgCl_2$ and 2 mM spermidine (lane 2), 6 mM $MgCl_2$ and 4 mM spermidine (lane 3), 10 mM $MgCl_2$ and 4 mM spermidine (lane 4) and 10 mM $MgCl_2$ and 1.3 mM spermidine (lane 5).

In a range between pH 7.5 and 9.0 at least 80 % of the enzyme activity was detected (Figure III.m (a)). Comparable to results for SP6 and T7 RNAP (Butler & Chamberlin, 1982; Chamberlin & Ring, 1973; Melton *et al.*, 1984), a pH value of 7.9 was suited best for *in vitro* transcription using K1E polymerase. Low NaCl concentrations (0 - 30 mM) were shown to have no influence on RNA synthesis (Figure III.m (b)). Salt concentrations above 30 mM led to a decrease in RNA formation. For example, RNAP activity was reduced by 50 %, when the reaction buffer was supplemented with 70 mM NaCl. Like other phage derived RNAPs (Butler & Chamberlin, 1982; Chamberlin & Ring, 1973), the enzyme of K1E phage shows an essential demand for

Mg^{2+} ions. No polymerase activity was detected without addition of $MgCl_2$ to the *in vitro* transcription buffer (Figure III.m (c)). Furthermore, a distinct activity optimum was detected for 10 mM $MgCl_2$. Moreover, it was shown, that addition of 23 mM of EDTA to the reaction mixture completely abolished RNAP activity. The K1E RNAP enzyme of the tested preparations did not show the need for DTT addition to the reaction buffer. However, Butler and Chamberlin (Butler & Chamberlin, 1982) described, that the need for thiol-reducing agents like DTT in *in vitro* transcriptions using SP6 RNAP varied between different enzyme preparations, storage conditions, exposure to air and other not well defined factors. Consequently, since no negative effect of DTT was detected 10 mM DTT were kept in the *in vitro* transcription buffer. Finally, buffers were tested containing increasing spermidine concentrations. Similar to SP6 and T7 RNAP (Butler & Chamberlin, 1982; Chamberlin & Ring, 1973), the K1E RNA polymerase was highly stimulated by spermidine and showed an optimum at 4 mM (Figure III.m (d)). Without this additive only 16 % of the maximum activity remained. As shown above, application of the various tested components in the given concentrations facilitated maximum K1E RNAP activities when employed individually. Surprisingly, the combination of 10 mM $MgCl_2$ with 4 mM spermidine did not lead to further enhanced, but to significantly decreased activities of K1E RNAP (Figure III.m (e)). Thus, further screenings for an optimized buffer were necessary. Finally, the buffer enabling highest RNA yields was found to contain a combination of 10 mM $MgCl_2$ and 1.3 mM spermidine. This adapted buffer increased RNA yields of *in vitro* transcriptions employing K1E RNA polymerase by more than 40 % compared to the initial buffer (Figure III.m (e)). Further RNA *in vitro* synthesis experiments were carried out using this optimized K1E RNAP buffer.

In order to determine the temperature optimum for the K1E RNAP, enzyme activities were measured in assays performed at different temperatures between 15 °C to 55 °C. A sharp optimum of K1E RNAP activity between 37 and 40 °C was observed (Figure III.n (a)). The half life time of the K1E RNAP at 50 °C was determined to be 16.5 s. For K1E RNAP stored in buffer B (20 mM $NaPO_4$, 100 mM NaCl, 1 mM EDTA, 10 mM DTT, 0.1 % Triton X-100, pH 7.7) supplemented with 50 % glycerol, a half life time of 87 h at 25 °C was calculated.

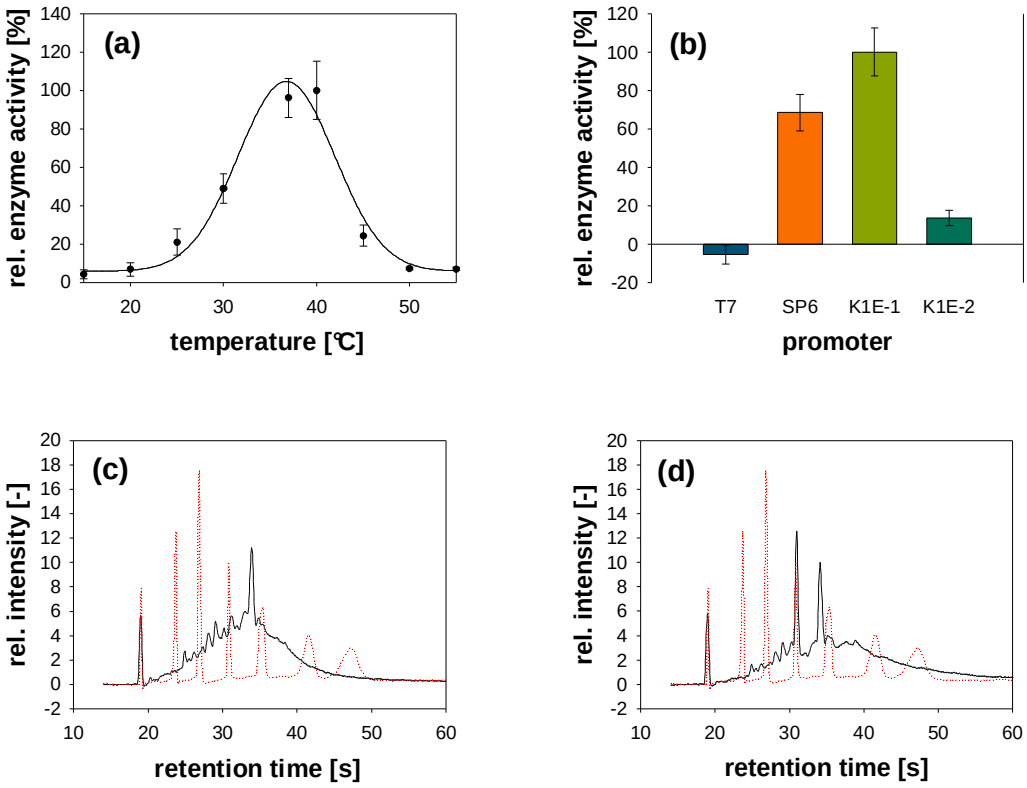

Figure III.n: Characterization of K1E RNA polymerase.

In *in vitro* transcription experiments the K1E RNAP was characterized with regard to **(a)** its temperature optimum, **(b)** its promoter and **(c)** and **(d)** its terminator preference. For this purpose, the optimized buffer conditions for K1E RNAP were applied as outlined in section II.6.11. In **(a)** and **(b)** maximum RNA yields correspond to 100 % relative enzyme activity. **(a)** Enzyme reactions were performed for 30 minutes at the indicated temperature. **(b)** Comparable plasmids were used as template differing in the indicated promoter sites used to initiate *in vitro* transcription. Panels **(c)** and **(d)** show electropherograms of RNA generated by *in vitro* syntheses using template plasmids **(d)** with or **(c)** without predicted K1E RNAP transcription terminator. The size standard is displayed as dotted red line. Its peaks represent RNAs of 25, 200, 500, 1,000, 2,000, 4,000 and 6,000 nucleotides in length.

III.3.4 Promoter specificity and terminator requirement of K1E RNA polymerase

Next, the promoter specificity of K1E RNAP was analyzed. For this purpose, plasmids containing either the T7 (taatacgactcactataggg), the SP6, the K1E-1 or the K1E-2 promoter were used as templates for *in vitro* transcriptions. Figure III.n (b)

shows that both K1E promoters (K1E-1 and K1E-2) facilitated transcription by K1E RNA polymerase. However, only 13 % of the *gfp*-mRNA was produced when using the plasmid harboring the K1E-2 promoter compared to the K1E-1 promoter carrying template. The K1E-2 sequence contains an adenine instead of the conserved guanine residue at the +1 position of the promoter. This result indicated that, similar to SP6 and T7 RNAP, also the K1E RNAP requires this guanine residue at +1 for significant transcriptional activity (Milligan *et al.*, 1987; Stump & Hall, 1993).

Like SP6 RNAP (Butler & Chamberlin, 1982), the K1E polymerase did not show any *gfp*-mRNA production from the T7 promoter. *Vice versa* T7 RNAP was not able to initiate transcription neither at the K1E-1 nor the K1E-2 promoter, indicating clear-cut promoter specificity (data not shown). Nevertheless, the SP6 promoter was recognized efficiently by K1E RNAP and facilitated the production of nearly 70 % of RNA compared to the most efficient promoter K1E-1. In agreement, SP6 RNAP (Promega Corporation) promoted RNA formation from both analyzed K1E promoters. However, RNA formation by SP6 RNAP was significantly less when employing one of the K1E promoters (K1E-1: 21 % and K1E-2: 4 %) instead of the appropriate SP6 pendant.

Finally, a functional transcription terminator was searched. For this purpose, a nucleotide BLAST with the K1E genome was performed using the predicted transcriptional terminator downstream of *orf* 31 of the SP6 genome as template (Scholl *et al.*, 2004). A DNA sequence showing 100 % identity was found at a comparable position within the genome of enterobacteriophage K1E. This DNA sequence (ccctatctaccttgcgtaggtagggttctttt) was tested for its ability to terminate transcription of K1E RNAP *in vitro*. To visualize the effectiveness of the phage RNAP transcriptional terminator, two comparable plasmids carrying the K1E-1 promoter upstream of the model gene *gfp*, one with (Figure III.n (d)) and the other without putative terminator (Figure III.n (c)), were subjected to *in vitro* transcription assays. A defined peak indicating a RNA with a length of approximately 1,000 nucleotides was visible in the electrophoretogram of the RNA fraction derived from the transcription of the terminator-containing *gfp* construct (Figure III.n (d)). Nevertheless, still significant read-through transcripts were detected. Although high amounts of mRNA were formed during *in vitro* transcription of the *gfp* gene without transcription terminator, no distinct peak of similar size was visible (Figure III.n (c)). This clearly indicates the functionality of the annotated K1E transcription terminator *in vitro*.

III.3.5 Large scale *in vitro* RNA production using K1E RNA polymerase

After determining appropriate assay conditions for *in vitro* transcriptions using K1E RNAP, a system for synthesis of large RNA amounts was established. For this purpose, the plasmid pK1Eivt was created by inserting the most effective promoter (K1E-1) into a pUC18 plasmid. The pUC18 basis was chosen due to its high-copy number mediating origin of replication for *E. coli*, which enabled the preparation of high amounts of template DNA. Another feature inherited from the pUC18 plasmid was the blue / white screening for successful cloning. The K1E-1 promoter was inserted upstream of the plasmid's multiple cloning site and the *lacZ* gene. Thereby, the *lacZ* fragment remained unchanged and the produced protein maintained its full α-complementation ability. Thus, positive *E. coli* colonies harboring a DNA insert within the *mcs* develop white colonies, while cells harboring plasmids without inserts form a blue dye under the appropriate conditions. This functionallity was proven during the construction of the test plasmid pK1Eivt-NIAysic1. Here, the coding region for an exogenous RNA control derived from the yeast genome, applicable in microarray or RT-PCR applications (Carter *et al.*, 2005) served as model and was inserted into pK1Eivt. Large scale synthesis of RNA via *in vitro* transcription using *Hind*III-linearized pK1Eivt-NIAysic1 plasmid as template yielded more than 5 µg purified RNA per µg plasmid DNA. Furthermore, Bioanalyzer data revealed the presence of only one RNA species with a distinct size of approximately 700 nucleotides corresponding to the desired target (data not shown).

III.3.6 Development of K1E and SP6 RNA polymerase driven recombinant protein production systems for *Bacillus megaterium*

New *in vivo* protein production systems for *B. megaterium* were developed employing K1E and SP6 RNAP for target gene transcription. For this purpose, the corresponding RNAP genes *k1ep* and *sp6p* were cloned into a plasmid under control of the homologous *B. megaterium* derived xylose-inducible promoter P_{xylA} (pK1E-RNAP, pSP6-RNAP). A second compatible, stable maintained vector was

equipped with either the commonly used SP6 promoter or the K1E-1 promoter and appropriate terminators. A ribosome binding site (*rbs*) adapted to *B. megaterium* was employed since it has been shown before to significantly enhance the translational efficiency (sections III.1.2 and III.1.3).

III.3.6.1 Intracellular production of Gfp by the new phage RNA polymerase driven systems

To evaluate the intracellular protein production of these phage RNAP driven systems, the model gene *gfp* was cloned under control of the SP6 and K1E phage RNAP promoters. First of all, *B. megaterium* MS941 cells were transformed with each of the phage RNAP encoding vectors individually. Then, transformed cells were protoplasted and transformed again. This time the second plasmid which carried *gfp* under control of one of the phage RNAP promoters was used for the transformation. Finally, Gfp production experiments were carried out with these two vector-harboring *B. megaterium* cells. Besides the natural combinations of SP6 RNAP (SP6 promoter) and K1E RNAP (K1E-1 promoter), the combination of K1E RNAP with the SP6 promoter was tested, which worked quite well *in vitro*.

It was shown that both the SP6 and the K1E RNAP were applicable for protein production in *B. megaterium* (Figure III.o).

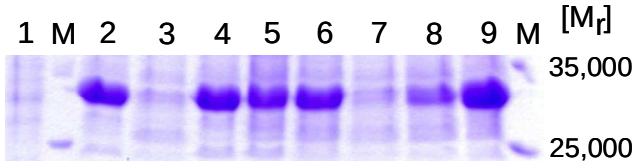

Figure III.o: Phage RNAP-dependent cytosolic production of Gfp evaluated by SDS-PAGE.

B. megaterium transformed with two plasmids, one comprising the phage RNAP gene under control of P_{xylA}, the other encoding Gfp under control of a phage RNAP promoter were employed in protein production experiments. Cultivations were performed in LB medium supplemented with appropriate antibiotics. At OD_{578nm} = 0.4 production of Gfp was induced by addition of 0.5 % (w/v) xylose. Samples of soluble cytosolic protein fraction were taken before (lane 1), 3 h after (lanes 2 - 5) and 7.5 h after (lanes 6 - 9) addition of the inducer of heterologous gene expression. Gfp amounts produced by *B. megaterium* cells harboring the T7 system (Gamer *et al.*, 2009) (lanes 2 and 6) were compared to those formed by bacteria employing one of the newly developed systems. The systems comprised the SP6 RNAP and SP6 promoter (lanes 3 and 7), the K1E RNAP and its cognate promoter K1E-1 (lanes 4 and 8) or the K1E RNAP and the SP6 promoter (lanes 1, 5 and 9). The unstained molecular weight marker (Fermentas) was used as size standard (M).

Maximum Gfp amounts produced by cells employing the K1E RNAP in combination with its natural promoter (K1E system) were nearly 9-fold higher (38.1 mg g_{CDW}^{-1}) compared to the amounts formed by bacteria employing the SP6 polymerase and its appropriate promoter (4.3 mg g_{CDW}^{-1}) (Figure III.p).

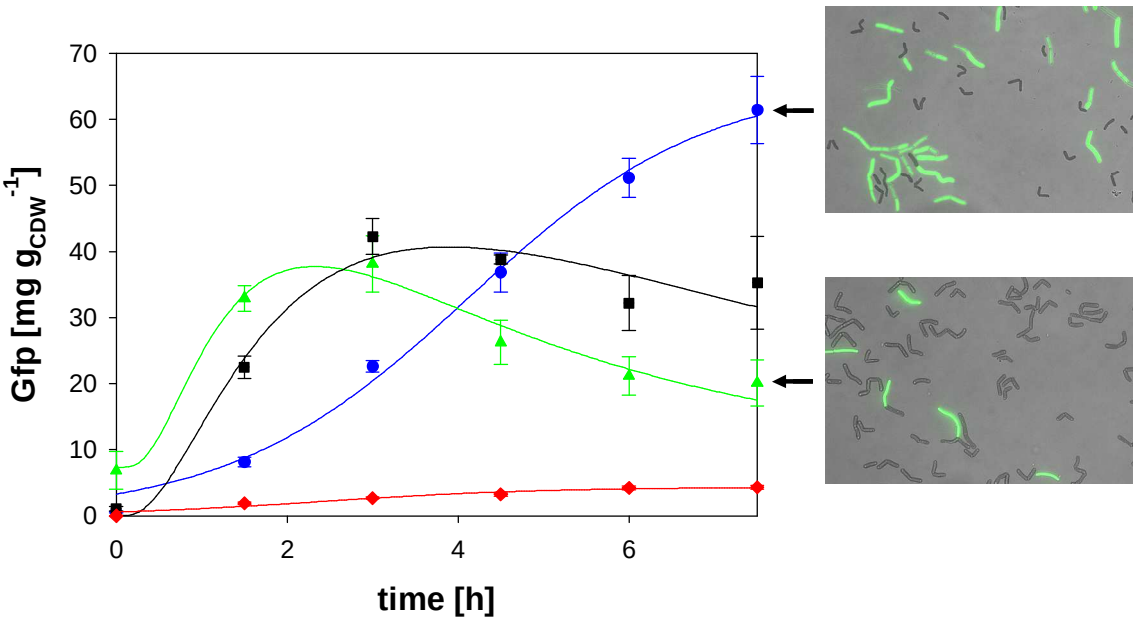

Figure III.p: Gfp production of *Bacillus megaterium* harboring phage RNAP-based expression systems.

B. megaterium MS941 cells were transformed with two plasmids. One was carrying the gene encoding the phage RNAP under control of P_{xylA}, the second was harboring the *gfp* gene under control of the phage RNAP promoter. All cultivations were performed at 37 °C in LB medium supplemented with the appropriate antibiotics. Production of Gfp was induced with 0.5 % (w/v) xylose at OD_{578nm} = 0.4. Samples were taken at the indicated times after xylose addition. Gfp amounts were calculated using data from fluorescence measurement of whole bacterial cells as described in section II.6.9. Cultivated *B. megaterium* cells harbored the K1E RNAP and the SP6 promoter (blue), the K1E RNAP and its cognate consensus promoter K1E-1 (green), the SP6 RNAP and the SP6 promoter (red), and the T7 system developed by Gamer *et al.* (2009) (black) which served as reference. On the right-hand side, fluorescence microscopic pictures of *B. megaterium* cells are shown. They refer to the cultures and times indicated by the arrows. Gfp producing cells appear bright green due to their fluorescence. To increase visibility on the picture, subsequently the contrast of non-producing cells (dark grey) was manually reworked.

Since both enzymes revealed quite comparable biochemical characteristics during their *in vitro* characterization (section III.3.3) (Butler & Chamberlin, 1982), this was a

quite surprising result. Further, comparison of the codon adaptation indices (CAIs) of the corresponding RNAP genes using the software “JCat” (http://www.jcat.de/) (Grote *et al.*, 2005) revealed an even better codon adaptation of the SP6 RNAP gene to *B. megaterium* codon usage (0.42) compared to the K1E RNAP gene (0.36). A lower CAI of a DNA sequence stands for a higher frequency of codons, which are of rare use in the genes of the organism. For the production of a hydrolase from *T. fusca* in *B. megaterium* it was shown, that the use of rare codons in the coding sequence leads to a slowdown of the translational process or even to complete abortion of translation (Yang *et al.*, 2007). Thus, the transcription of the RNAP gene did not seem to be the limiting step of the SP6 RNAP driven expression system. It seemed more likely, that either the SP6 RNAP was degraded faster within the host *B. megaterium* or it was *in vivo* less active compared to the K1E enzyme.

When comparing the K1E system and the reference T7 system for *B. megaterium* (Gamer *et al.*, 2009), the temporal progress of protein production was quite similar. Both production processes showed a strong increase of Gfp over the first hours after addition of the gene expression inducer xylose. Furthermore, a comparable production maximum of heterologous protein was reached (T7 system: 42.3 mg g_{CDW}^{-1} and K1E system: 38.1 mg g_{CDW}^{-1}). Compared to the T7 system, a slightly faster decrease in the Gfp amount per cell dry weight was detected for the K1E system. Cells harboring the K1E system already showed significant Gfp amounts (6.9 mg g_{CDW}^{-1}) prior to the addition of the gene expression inducer xylose.

Interestingly, *B. megaterium* cells harboring the plasmid encoded K1E RNAP combined with *gfp* under control of the *in vitro* less effective SP6 promoter showed the maximal Gfp production (61.4 mg g_{CDW}^{-1}) *in vivo*. Cells of these cultures showed continuously increasing intracellular Gfp amounts until 7.5 h after induction of heterologous gene expression. However, the slope of the production curve was significantly lower than in the initial production phase of *B. megaterium* employing either the T7 system or the K1E system. Nevertheless, application of different promoters (K1E-1 and SP6) in the K1E RNAP driven protein production systems revealed, that the less effective SP6 promoter led to a slower, but longer lasting increase of target protein, finally yielding 60 % more Gfp (61.4 mg g_{CDW}^{-1}) when compared to the system employing the consensus K1E-1 promoter (38.1 mg g_{CDW}^{-1}). Fluorescence microscopy (Figure III.p) unveiled that bacterial cultures employing the SP6 promoter driven K1E RNAP system showed a non-productive subpopulation of

minor extent. During the protein production process the percentage of non-producing cells remained nearly constant. A similar culture heterogeneity was already observed in protein producing *B. megaterium* cultures before (Biedendieck *et al.*, 2007c). However, a significant higher portion of non-productive (non-fluorescent) cells accumulated over time in cultures using the K1E-1 promoter controlled K1E RNAP system. Patkar *et al.* (2002) reported about a comparable phenomenon when applying a T7 RNAP driven gene expression system in *E. coli* cells. They proposed that due to reduced growth rates of the highly productive *E. coli* cells, plasmid-free cells devoid of target gene expression began to outgrow the culture (Patkar *et al.*, 2002). Alternatively, an individual stress response induced by the metabolic burden of recombinant protein production might have led to bistability like it was found under certain conditions for other bacterial cultures including *B. subtilis* (Dubnau & Losick, 2006). For the elucidation of the molecular mechanism for this culture heterogeneity phenomenon further work is needed.

III.3.6.2 Production and secretion of the model protein Tfh by the K1E-based production systems

In 2009, Gamer *et al.* reported about inefficient secretion of the model protein LevΔ773 when comparing the *B. megaterium* T7 protein production system with the P_{xylA}-based one. Analysis of subcellular fractions revealed an intracellular accumulation of unprocessed target protein. This indicated a limitation in the secretion process. Further, accumulation of mature protein suggested a second drawback in diffusion through the cell wall. This might be caused by the high relative molecular mass of LevΔ773 of 110,000 (Biedendieck *et al.*, 2007a). As a consequence, in this work the model enzyme Tfh (M_r = 28,000) was employed. Further, in accordance with the results of section III.3.6.1, it was suggested that the major drawback in secretory production of recombinant proteins using the T7-RNAP system might be that the accumulation of target protein occurred to rapidly and resulted in the overgrowth of the culture by non-productive bacterial cells. Thus, it was suggested that the K1E-RNAP system employing the SP6 promoter might be much better suited for secretory protein production. For this system the culture

heterogeneity is much less pronounced than for the T7- or the K1E-RNAP system (employing its cognate promoter) (Figure III.p).

For verification of this suggestion, the target gene *tfh-his*$_6$ fused to the *sp*$_{yocH}$ was cloned into the plasmids carrying the T7, SP6 or K1E-1 RNAP promoter. *B. megaterium* MS941 cells with plasmids harboring the K1E-RNAP or T7-RNAP were transformed with these plasmids appropriately. Finally, three phage RNAP-based systems were employed for Tfh production and secretion. These were, the T7 system (Gamer *et al.*, 2009), the system employing K1E RNAP and its natural promoter K1E-1 and the K1E system using the SP6 promoter for initiation of transcription. The strain *B. megaterium* MS941 pSSBm97 (P$_{xylA}$-(-35$^+$+*rbs*$^+$)-*sp*$_{yocH}$-*tfh-his*$_6$) employing a hp-plasmid and the SP$_{YocH}$ for secretion of Tfh was used as reference.

Data showed similar results for Tfh as found by Gamer *et al.* (2009) for the production and export of heterologous LevΔ773 (Figure III.q). The T7 system yielded significantly less secreted heterologous Tfh-His$_6$ (max. 1,500 U mL^{-1}) compared to the P$_{xylA}$-based system (max. 6,500 U mL^{-1}) (Figure III.r).

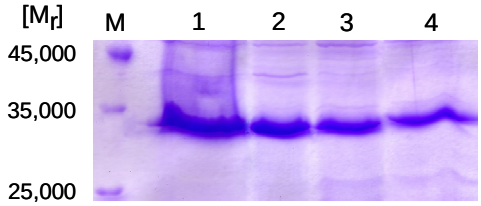

Figure III.q: SDS-PAGE analysis of Tfh containing supernatants of *Bacillus megaterium* cultures harboring phage RNAP-based expression systems.

B. megaterium MS941 cells with plasmids encoding phage T7 or K1E RNAP were transformed with appropriate plasmids encoding SP$_{yocH}$-Tfh-His$_6$ fusion pre-protein under control of a phage RNAP promoter. The *B. megaterium* MS941 strain carrying the corresponding hp-plasmid pSSBm97 (P$_{xylA}$-(-35$^+$+*rbs*$^+$)-*sp*$_{yocH}$-*tfh-his*$_6$) was employed as a reference. Cultivations were performed aerobically in LB medium supplemented with the appropriate antibiotics at 37 °C. Heterologous production of Tfh was induced with 0.5 % (w/v) xylose at OD$_{578nm}$ = 0.4. Samples were taken 7.5 h after xylose addition. Proteins of 1.5 mL cell-free supernatant of *B. megaterium* cultures were precipitated with ammonium sulfate. Supernatant samples were taken of the reference strain (lane 1), of bacterial cultures employing K1E RNAP in combination with the SP6 promoter (lane 2), of cells harboring the K1E RNAP and its cognate K1E-1 promoter (lane 3) and of *B. megaterium* using the T7 system (lane 4) (Gamer *et al.*, 2009). Proteins were separated on 12 % SDS PAGE gels and visualized by Coomassie Brilliant Blue staining. The unstained protein molecular weight marker (Fermentas) was used as size standard (M).

When K1E-RNAP is employed with its cognate promoter (K1E-1), the system worked just little more effective (max. 1,900 U mL^{-1}) than the T7 system (max. 1,500 U mL^{-1}). However, non of these systems was able to facilitate the export of more than 29 % of the Tfh-His_6 produced and secreted by the P_{xylA}-based system. When applying the K1E-RNAP in combination with the SP6-promoter for *tfh* transcription volumetric protein yields (max. 3,000 U mL^{-1}) could be increased by more than 58 % compared the other phage RNAP-based systems. However, this is still just 46 % of the Tfh amount secreted by the P_{xylA}-based system. This was astonishing, since the temporal progress of Tfh production inside the cell when using the K1E system employing the SP6 promoter is quite comparable to the one of the P_{xylA} system as to be seen from data of Gfp formation (compare Biedendieck *et al.* (2007c) and section III.3.6.1).

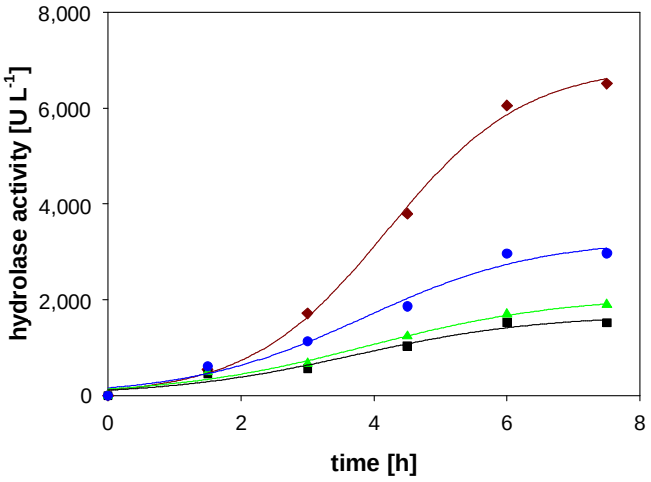

Figure III.r: Production and secretion of Tfh applying phage RNAP-based production systems in *Bacillus megaterium*

B. megaterium transformed with two plasmids, one comprising the phage RNAP gene, the other encoding the SP_{YocH}-Tfh-His_6 pre-protein fusion under control of a phage RNAP promoter were employed in protein production experiments. *B. megaterium* MS941 cells harboring the corresponding hp-plasmid pSSBm97 (P_{xylA}-(-35^++*rbs*$^+$)-sp_{yocH}-*tfh*-his_6) were employed as reference. All strains were cultivated aerobically in LB medium supplemented with the appropriate antibiotics at 37 °C. Heterologous production and secretion of Tfh was induced with 0.5 % (w/v) xylose at OD_{578nm} = 0.4. Samples were taken every 1.5 h after xylose addition. Hydrolase activities were determined photometrically at 30 °C by the *p*-nitrophenylpalmitate (*p*NPP) method as described in section II.6.10.. Supernatants of *B. megaterium* cells harboring the K1E RNAP and the SP6 promoter (blue), the K1E RNAP and its cognate consensus promoter K1E-1 (green), the T7 system (black) or the reference system (dark red) were analyzed.

Thus, it was suggested that a hampering of the secretion process due to an accumulation of pre-protein to be exported would be significantly reduced for this phage RNAP-based system. Due to the much lower relative molecular mass, diffusion limitation as found for the LevΔ773 should be significantly reduced for the secretion of Tfh-His_6. Anyway, this limitation is supposed to be present in all cells regardless of whether they carry the P_{xylA} or a phage RNAP-based expression system. Thus, there must be another significant limitation hampering the secretion of cells employing one of the phage RNAP-based production systems. Maybe it is linked to the higher metabolic burden, due to the maintenance of two plasmids, the additional production of the RNAP enzyme and the struggle against two antibiotics within the culture medium.

IV. Summary

The established *B. megaterium* protein production system based on the xylose-inducible promoter (P_{xylA}) (Biedendieck *et al.*, 2007c) was systematically optimized. Multiple changes in basic promoter elements such as the -10 and -35 region and the ribosome-binding site resulted in an 18-fold increase of protein production. This new high performance (hp)-system enabled the production of 82.5 mg g_{CDW}^{-1} of the model protein Gfp. Furthermore, signal peptides for Sec-dependent protein secretion were predicted *in silico* using the *B. megaterium* genome sequence. Subsequently, the leader peptides of Vpr, NprM, YngK, YocH and a computationally designed peptide were experimentally validated. Highest yields were obtained using *B. megaterium* carrying a hp-expression plasmid encoding SP_{YocH} for export of Tfh (7,200 U L^{-1} or 7.7 mg L^{-1}). Nevertheless, observed principle limitations in protein export might be related to components of the Sec-dependent protein transport system.

Further, a novel P_{xylA}-based gene expression plasmid was developed for use in *B. megaterium*. It is equipped with an “on command” copy number amplification system for *E. coli*, designed to decrease undesired basal expression of P_{xylA} during the *E. coli* based cloning procedure. The functionality of this system in *E. coli* was demonstrated. In *B. megaterium*, 5.2 mg g_{CDW}^{-1} of Gfp and 1,800 U L^{-1} (equals 2.0 mg L^{-1}) of TfH were produced and secreted, respectively.

The predicted RNA polymerase (RNAP) of *E. coli* phage K1E was produced recombinantly, purified to apparent homogeneity and employed for *in vitro* RNA synthesis. Optimal assay conditions (pH 8, 37 °C, 10 mM $MgCl_2$ and 1.3 mM spermidine) were established. The corresponding promoter regions were identified on the phage genome and summarized in a sequence logo. Surprisingly, next to K1E promoters also the consensus SP6 promoter was recognized efficiently *in vitro* by K1E RNAP. Based on these results, a system for high yield *in vitro* RNA synthesis (5 µg purified RNA per µg template DNA) using K1E RNAP was established. Finally, *in vivo* protein production systems for *B. megaterium* were established on K1E phage RNAP transcription. Up to 61.4 mg g_{CDW}^{-1} of the reporter protein Gfp were produced and up to 3,000 U L^{-1} (equals 3.2 mg L^{-1}) of extracellular Tfh were formed. This thesis established novel, now commercially available high yield protein production and export systems for *B. megaterium*.

V. Outlook

Since 2010, the complete genome sequence of *B. megaterium* DSM319 is publicly available at the National Center for Biotechnology Information (NCBI). Now, this data enables systems biotechnological studies of the protein production and secretion process in the bacterial host *B. megaterium*. In addition to the results described in this thesis, appropriate methods for transcriptome, metabolome and proteome analysis of *B. megaterium* were established during this work. Data sets of bacterial cultures producing heterologous proteins derived with these technologies might help to pinpoint bottlenecks during cultivation and secretion. Subsequently, targeted genetic engineering might promote a further increased productivity.

Due to the newly developed and optimized recombinant gene expression tools for *B. megaterium*, significant increase in intracellular production of the model protein Gfp was observed during this work. Next, these results have to be consolidated by production of other performance proteins or enzymes of industrial and scientific importance to strengthen *B. megaterium*'s position as arinsing alternative host system.

However, the major advantage of *B. megaterium*, its high natural secretory capacity, is still disappointing with regard to the secretion of recombinant heterologous proteins. Neither higher accumulation of exo-pre-proteins inside the cell, nor the use of novel signal peptides could significantly enhance the amount of proteins secreted via the Sec-dependent pathway. Thus, further optimization of the *B. megaterium* Sec-system is advised. It would be possible to specifically overexpress components of the Sec-secretion apparatus to enhance protein transport. However, since the transport mechanism is quite complex and not well known especially with respect to heterologous protein production and secretion, a random approach might be even more promising.

Finally, the genetic accessibility of *B. megaterium* has still to be improved. This could significantly simplify genetic engineering of the host chromosome and would customize *B. megaterium* for purposes like *e.g.* enzyme variant screenings.

VI. References

Ash, C., Farrow, J. A. E., Wallbanks, S. & Collins, M. D. (1991). Phylogenetic heterogeneity of the genus *Bacillus* revealed by comparative analysis of small-subunit-ribosomal RNA sequences. *Lett Appl Microbiol* **13**, 202-206.

Atiyeh, C. (2010). Bio-based Chemicals: Sustainable, Economical and Marketable. *The World Congress on Biotechnology & Bioprocessing*. Washington D.C.

Beller, M., Böhland, T., Demtröder, D. & other authors (2010). Positionspapier: Rohstoffbasis im Wandel. DECHEMA. Frankfurt. http://www.dechema.de/.

Biedendieck, R. (2006). *Bacillus megaterium*: Versatile tools for production, secretion and purification of recombinant proteins. PhD-Thesis. *Institute of Microbiology*, Technische Universität Braunschweig, Braunschweig.

Biedendieck, R., Beine, R., Gamer, M., Jordan, E., Buchholz, K., Seibel, J., Dijkhuizen, L., Malten, M. & Jahn, D. (2007a). Export, purification, and activities of affinity tagged *Lactobacillus reuteri* levansucrase produced by *Bacillus megaterium*. *Appl Microbiol Biotechnol* **74**, 1062-1073.

Biedendieck, R., Gamer, M., Jaensch, L., Meyer, S., Rohde, M., Deckwer, W. D. & Jahn, D. (2007b). A sucrose-inducible promoter system for the intra- and extracellular protein production in *Bacillus megaterium*. *J Biotechnol* **132**, 426-430.

Biedendieck, R., Yang, Y., Deckwer, W. D., Malten, M. & Jahn, D. (2007c). Plasmid system for the intracellular production and purification of affinity-tagged proteins in *Bacillus megaterium*. *Biotechnol Bioeng* **96**, 525-537.

Biedendieck, R., Malten, M., Barg, H., Bunk, B., Martens, J.-H., Deery, E., Leech, H., Warren, M. J. & Jahn, D. (2010). Metabolic engineering of cobalamin (vitamin B_{12}) production in *Bacillus megaterium*. *Microb Biotechnol* **3**, 24-37.

Brockmeier, U., Caspers, M., Freudl, R., Jockwer, A., Noll, T. & Eggert, T. (2006). Systematic screening of all signal peptides from *Bacillus subtilis*: a powerful strategy in optimizing heterologous protein secretion in Gram-positive bacteria. *J Mol Biol* **362**, 393-402.

Bunk, B. (2010). Comparative and Functional Genomics of *Bacillus megaterium* DSM319. PhD-Thesis. *Institute of Microbiology*, Technische Universität Braunschweig, Braunschweig.

Butler, E. T. & Chamberlin, M. J. (1982). Bacteriophage SP6-specific RNA polymerase. I. Isolation and characterization of the enzyme. *J Biol Chem* **257**, 5772-5778.

Carter, M. G., Sharov, A. A., VanBuren, V., Dudekula, D. B., Carmack, C. E., Nelson, C. & Ko, M. S. (2005). Transcript copy number estimation using a mouse whole-genome oligonucleotide microarray. *Genome Biol* **6**, R61.

Chakraborty, P. R., Salvo, R. A., Majumder, H. K. & Maitra, U. (1977). Further characterization of bacteriophage T3-induced ribonucleic acid polymerase. Studies on the size of *in vitro* transcripts and interaction of T3 RNA polymerase with T3 DNA. *J Biol Chem* **252**, 6485-6493.

Chamberlin, M. & Ring, J. (1973). Characterization of T7-specific ribonucleic acid polymerase. 1. General properties of the enzymatic reaction and the template specificity of the enzyme. *J Biol Chem* **248**, 2235-2244.

Chamberlin, M. & Ryan, T. (1982). Bacteriophage DNA-dependent RNA polymerases. In *The Enzymes*, pp. 87-109. Edited by P. D. Boyer. New York: Academic Press.

Chen, Z. & Schneider, T. D. (2005). Information theory based T7-like promoter models: classification of bacteriophages and differential evolution of promoters and their polymerases. *Nucleic Acids Res* **33**, 6172-6187.

Conrad, B., Savchenko, R. S., Breves, R. & Hofemeister, J. (1996). A T7 promoter-specific, inducible protein expression system for *Bacillus subtilis*. *Mol Gen Genet* **250**, 230-236.

Cormack, B. P., Valdivia, R. H. & Falkow, S. (1996). FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* **173**, 33-38.

Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome Res* **14**, 1188-1190.

De Bary, A. (1884). *Vergleichende Morphologie und Biologie der Pilze, Mycetozoen und Bakterien*. Leipzig, Germany: Wilhelm Engelmann.

deHaseth, P. L., Zupancic, M. L. & Record, M. T., Jr. (1998). RNA polymerase-promoter interactions: the comings and goings of RNA polymerase. *J Bacteriol* **180**, 3019-3025.

Delcher, A. L., Harmon, D., Kasif, S., White, O. & Salzberg, S. L. (1999). Improved microbial gene identification with GLIMMER. *Nucleic Acids Res* **27**, 4636-4641.

Dresler, K., van den Heuvel, J., Muller, R. J. & Deckwer, W. D. (2006). Production of a recombinant polyester-cleaving hydrolase from *Thermobifida fusca* in *Escherichia coli*. *Bioprocess Biosyst Eng* **29**, 169-183.

Dubnau, D. & Losick, R. (2006). Bistability in bacteria. *Mol Microbiol* **61**, 564-572.

Durrschmid, K., Reischer, H., Schmidt-Heck, W., Hrebicek, T., Guthke, R., Rizzi, A. & Bayer, K. (2008). Monitoring of transcriptome and proteome profiles to

investigate the cellular response of *E. coli* towards recombinant protein expression under defined chemostat conditions. *J Biotechnol* **135**, 34-44.

Ellinger, T., Behnke, D., Bujard, H. & Gralla, J. D. (1994). Stalling of *Escherichia coli* RNA polymerase in the +6 to +12 region *in vivo* is associated with tight binding to consensus promoter elements. *J Mol Biol* **239**, 455-465.

Emanuelsson, O., Brunak, S., von Heijne, G. & Nielsen, H. (2007). Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc* **2**, 953-971.

Freudl, R. (1992). Protein secretion in gram-positive bacteria. *J Biotechnol* **23**, 231-240.

Fuerst, T. R., Niles, E. G., Studier, F. W. & Moss, B. (1986). Eukaryotic transient-expression system based on recombinant vaccinia virus that synthesizes bacteriophage T7 RNA polymerase. *Proc Natl Acad Sci U S A* **83**, 8122-8126.

Gamer, M., Fröde, D., Biedendieck, R., Stammen, S. & Jahn, D. (2009). A T7 RNA polymerase-dependent gene expression system for *Bacillus megaterium*. *Appl Microbiol Biotechnol* **82**, 1195-1203.

Golomb, M. & Chamberlin, M. (1974). Characterization of T7-specific ribonucleic acid polymerase. IV. Resolution of the major *in vitro* transcripts by gel electrophoresis. *J Biol Chem* **249**, 2858-2863.

Grote, A., Hiller, K., Scheer, M., Münch, R., Nörtemann, B., Hempel, D. C. & Jahn, D. (2005). JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. *Nucleic Acids Res* **33**, W526-531.

Guzman, L. M., Belin, D., Carson, M. J. & Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter. *J Bacteriol* **177**, 4121-4130.

Heim, R., Prasher, D. C. & Tsien, R. Y. (1994). Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proc Natl Acad Sci U S A* **91**, 12501-12504.

Heinemann, I. U. (2007). Structure and function of enzymes involved in tetrapyrrole biosynthesis. PhD-Thesis. *Insitute of Microbiology*, Technische Universität Braunschweig, Braunschweig.

Hiller, K., Grote, A., Scheer, M., Münch, R. & Jahn, D. (2004). PrediSi: prediction of signal peptides and their cleavage positions. *Nucleic Acids Res* **32**, W375-379.

Hong, H. A., Huang, J. M., Khaneja, R., Hiep, L. V., Urdaci, M. C. & Cutting, S. M. (2008). The safety of *Bacillus subtilis* and *Bacillus indicus* as food probiotics. *J Appl Microbiol* **105**, 510-520.

Jorgensen, E. D., Durbin, R. K., Risman, S. S. & McAllister, W. T. (1991). Specific contacts between the bacteriophage T3, T7, and SP6 RNA polymerases and their promoters. *J Biol Chem* **266**, 645-651.

Keeling, P. (2010). An Array of Biorenewable Chemicals through Fermentation and Catalysis. *The World Congress on Biotechnology & Bioprocessing*. Washington D.C.

Kieselburg, M. K., Weickert, M. & Vary, P. S. (1984). Analysis of resident and transformant plasmids in *Bacillus megaterium*. *Bio/Technology* **2**, 254-259.

Kleeberg, I., Welzel, K., Vandenheuvel, J., Müller, R. J. & Deckwer, W. D. (2005). Characterization of a new extracellular hydrolase from *Thermobifida fusca* degrading aliphatic-aromatic copolyesters. *Biomacromolecules* **6**, 262-270.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

Lavigne, R., Seto, D., Mahadevan, P., Ackermann, H. W. & Kropinski, A. M. (2008). Unifying classical and molecular taxonomic classification: analysis of the Podoviridae using BLASTP-based tools. *Res Microbiol* **159**, 406-414.

Lobet, Y., Peacock, M. G. & Cieplak, W., Jr. (1989). Frame-shift mutation in the *lacZ* gene of certain commercially available pUC18 plasmids. *Nucleic Acids Res* **17**, 4897.

Malten, M., Hollmann, R., Deckwer, W. D. & Jahn, D. (2005). Production and secretion of recombinant *Leuconostoc mesenteroides* dextransucrase DsrS in *Bacillus megaterium*. *Biotechnol Bioeng* **89**, 206-218.

Malten, M., Biedendieck, R., Gamer, M., Drews, A. C., Stammen, S., Buchholz, K., Dijkhuizen, L. & Jahn, D. (2006). A *Bacillus megaterium* plasmid system for the production, export, and one-step purification of affinity-tagged heterologous levansucrase from growth medium. *Appl Environ Microbiol* **72**, 1677-1679.

McAllister, W. T. (1991). Plasmid for the overproduction of bacteriophage T3 RNA polymerase, transcription vectors that carry a promoter recognized by its polymerase, gene coding for T3 RNA polymerase and application of these plasmids. In *US Patent*. USA: University of Medicine and Dentistry of New Jersey.

Melton, D. A., Krieg, P. A., Rebagliati, M. R., Maniatis, T., Zinn, K. & Green, M. R. (1984). Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucleic Acids Res* **12**, 7035-7056.

Milligan, J. F., Groebe, D. R., Witherell, G. W. & Uhlenbeck, O. C. (1987). Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res* **15**, 8783-8798.

Müller, R. J., Schrader, H., Profe, J., Dresler, K. & Deckwer, W. D. (2005). Enzymatic degradation of poly(ethylene terephthalate): rapid hydrolyse using a

hydrolase from *T. fusca*. *Macromol Rapid Commun* **26**, 1400-1405.

Münch, R., Hiller, K., Grote, A., Scheer, M., Klein, J., Schobert, M. & Jahn, D. (2005). Virtual Footprint and PRODORIC: an integrative framework for regulon prediction in prokaryotes. *Bioinformatics* **21**, 4187-4189.

Panbangred, W., Weeradechapon, K., Udomvaraphant, S., Fujiyama, K. & Meevootisom, V. (2000). High expression of the penicillin G acylase gene (*pac*) from *Bacillus megaterium* UN1 in its own pac minus mutant. *J Appl Microbiol* **89**, 152-157.

Paschal, B. M., McReynolds, L. A., Noren, C. J. & Nichols, N. M. (2008). RNA polymerases. *Curr Protoc Mol Biol* **Chapter 3**, Unit3 8.

Patkar, A., Vijayasankaran, N., Urry, D. W. & Srienc, F. (2002). Flow cytometry as a useful tool for process development: rapid evaluation of expression systems. *J Biotechnol* **93**, 217-229.

Qiu, J. & Helmann, J. D. (1999). Adenines at -11, -9 and -8 play a key role in the binding of *Bacillus subtilis* Esigma(A) RNA polymerase to -10 region single-stranded DNA. *Nucleic Acids Res* **27**, 4541-4546.

Raux, E., Lanois, A., Warren, M. J., Rambach, A. & Thermes, C. (1998). Cobalamin (vitamin B_{12}) biosynthesis: identification and characterization of a *Bacillus megaterium cobI* operon. *Biochem J* **335 (Pt 1)**, 159-166.

Reischer, H., Schotola, I., Striedner, G., Potschacher, F. & Bayer, K. (2004). Evaluation of the GFP signal and its aptitude for novel on-line monitoring strategies of recombinant fermentation processes. *J Biotechnol* **108**, 115-125.

Riese, J. & Bachmann, R. (2004). Industrial Biotechnology: Turning the Potential into Profits. McKinsey & Company. http://www.mckinsey.com/.

Righetti, P. G., Gianazza, E., Gelfi, C. & Chairi, M. (1990). *Gel electrophoresis of proteins: a practical approach*, 2 edn. Oxford: Oxford Univeristy Press.

Roth, A. H. & Dersch, P. (2010). A novel expression system for intracellular production and purification of recombinant affinity-tagged proteins in *Aspergillus niger*. *Appl Microbiol Biotechnol* **86**, 659-670.

Ruiz, C., Blanco, A., Pastor, F. I. & Diaz, P. (2002). Analysis of *Bacillus megaterium* lipolytic system and cloning of LipA, a novel subfamily I.4 bacterial lipase. *FEMS Microbiol Lett* **217**, 263-267.

Rygus, T. & Hillen, W. (1991). Inducible high-level expression of heterologous genes in *Bacillus megaterium* using the regulatory elements of the xylose-utilization operon. *Appl Microbiol Biotechnol* **35**, 594-599.

Rygus, T., Scheler, A., Allmansberger, R. & Hillen, W. (1991). Molecular cloning, structure, promoters and regulatory elements for transcription of the *Bacillus megaterium* encoded regulon for xylose utilization. *Arch Microbiol* **155**, 535-542.

Sagawa, H., Ohshima, A. & Kato, I. (1996). A tightly regulated expression system in *Escherichia coli* with SP6 RNA polymerase. *Gene* **168**, 37-41.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (2001). *Molecular cloning, a laboratory manual*, 4 edn. New York: Cold Spring Harbor Laboratory Press.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74**, 5463-5467.

Schlax, P. J., Capp, M. W. & Record, M. T., Jr. (1995). Inhibition of transcription initiation by *lac* repressor. *J Mol Biol* **245**, 331-350.

Schneider, T. D. & Stephens, R. M. (1990). Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res* **18**, 6097-6100.

Scholl, D., Kieleczawa, J., Kemp, P., Rush, J., Richardson, C. C., Merril, C., Adhya, S. & Molineux, I. J. (2004). Genomic analysis of bacteriophages SP6 and K1-5, an estranged subgroup of the T7 supergroup. *J Mol Biol* **335**, 1151-1171.

Scholz, O., Thiel, A., Hillen, W. & Niederweis, M. (2000). Quantitative analysis of gene expression with an improved green fluorescent protein. p6. *Eur J Biochem* **267**, 1565-1570.

Schweizer, H. P., Klassen, T. R. & Hoang, T. (1996). Improved methods for gene analysis and expression in *Pseudomonas*. In *Molecular biology of pseudomonads*, pp. 229-237. Edited by T. Nakazawa, D. Haas & S. Silver. Washington, DC: ASM Press.

Shimomura, O., Johnson, F. H. & Saiga, Y. (1962). Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea. *J Cell Comp Physiol* **59**, 223-239.

Stammen, S., Müller, B. K., Korneli, C., Biedendieck, R., Gamer, M., Franco-Lara, E. & Jahn, D. (2010a). High-yield intra- and extracellular protein production using *Bacillus megaterium*. *Appl Environ Microbiol* **76**, 4037-4046.

Stammen, S., Schuller, F., Dietrich, S., Gamer, M., Biedendieck, R. & Jahn, D. (2010b). Application of *Escherichia coli* phage K1E DNA-dependent RNA polymerase for *in vitro* RNA synthesis and *in vivo* protein production in *Bacillus megaterium*. *Appl Microbiol Biotechnol* **88**, 529-539.

Studier, F. W., Davanloo, P., Rosenberg, A. H., Moffatt, B. A. & Dunn, J. J. (1999). Cloning and expression of the gene for bacteriophage T7 RNA polymerase. In *US Patent*. Upton, NY, USA: Brookhaven Science Associates LLC.

Stummeyer, K., Schwarzer, D., Claus, H., Vogel, U., Gerardy-Schahn, R. & Mühlenhoff, M. (2006). Evolution of bacteriophages infecting encapsulated bacteria: lessons from *Escherichia coli* K1-specific phages. *Mol Microbiol* **60**, 1123-1135.

Stump, W. T. & Hall, K. B. (1993). SP6 RNA polymerase efficiently synthesizes RNA from short double-stranded DNA templates. *Nucleic Acids Res* **21**, 5480-5484.

Tabor, S. & Richardson, C. C. (1985). A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc Natl Acad Sci U S A* **82**, 1074-1078.

Tjalsma, H., Bolhuis, A., Jongbloed, J. D., Bron, S. & van Dijl, J. M. (2000). Signal peptide-dependent protein transport in *Bacillus subtilis*: a genome-based survey of the secretome. *Microbiol Mol Biol Rev* **64**, 515-547.

Tjalsma, H., Antelmann, H., Jongbloed, J. D. & other authors (2004). Proteomics of protein secretion by *Bacillus subtilis*: separating the "secrets" of the secretome. *Microbiol Mol Biol Rev* **68**, 207-233.

Tsien, R. Y. (1998). The green fluorescent protein. *Annu Rev Biochem* **67**, 509-544.

Vary, P. S. (1994). Prime time for *Bacillus megaterium*. *Microbiology* **140 (Pt 5)**, 1001-1013.

Vary, P. S., Biedendieck, R., Fürch, T., Meinhardt, F., Rohde, M., Deckwer, W. D. & Jahn, D. (2007). *Bacillus megaterium* - from simple soil bacterium to industrial protein production host. *Appl Microbiol Biotechnol* **76**, 957-967.

von Tersch, M. A. & Robbins, H. L. (1990). Efficient cloning in *Bacillus megaterium*: comparison to *Bacillus subtilis* and *Escherichia coli* cloning hosts. *FEMS Microbiol Lett* **70**, 305-310.

Wang, W., Hollmann, R. & Deckwer, W. D. (2006). Comparative proteomic analysis of high cell density cultivations with two recombinant *Bacillus megaterium* strains for the production of a heterologous dextransucrase. *Proteome Sci* **4**, 19.

Wild, J., Hradecna, Z. & Szybalski, W. (2002). Conditionally amplifiable BACs: switching from single-copy to high-copy vectors and genomic clones. *Genome Res* **12**, 1434-1444.

Wittchen, K. D. & Meinhardt, F. (1995). Inactivation of the major extracellular protease from *Bacillus megaterium* DSM319 by gene replacement. *Appl Microbiol Biotechnol* **42**, 871-877.

Yamane, K., Bunai, K. & Kakeshita, H. (2004). Protein traffic for secretion and related machinery of *Bacillus subtilis*. *Biosci Biotechnol Biochem* **68**, 2007-2023.

Yang, Y., Malten, M., Grote, A., Jahn, D. & Deckwer, W. D. (2007). Codon optimized *Thermobifida fusca* hydrolase secreted by *Bacillus megaterium*. *Biotechnol Bioeng* **96**, 780-794.

VII. Danksagung

Zu allererst möchte ich Prof. Dr. Dieter Jahn danken mir die Möglichkeit eingeräumt zu haben, am Institut für Mikrobiologie der TU Braunschweig meine Dissertation zu diesem interessanten, anwendungsorientierten Thema anfertigen zu können. Die freie und selbstbestimmte Arbeitsweise habe ich sehr zu schätzen gelernt und danke für die dadurch erlangte Selbständigkeit in Forschung und Laborорganisation.
Des Weiteren danke ich Prof. Dr. Petra Dersch für die Übernahme des Korreferates und Prof. Dr. Christoph Wittmann für die Übernahme des Prüfungsvorsitzes.
Ein besonderer Dank gilt allen, die an der praktischen Durchführung von Versuchen beteiligt waren. Speziell möchte ich dabei Franziska Schuller und Katrin Müller erwähnen, die durch ihre Diplom- bzw. Studienarbeit einiges zu dieser Arbeit beitragen konnten. Auch sei allen Laborpraktikanten (Heinrich Schlums, Johannes Schwerk, Janika Viereck und Sylvia Dietrich) gedankt, die stets interessiert und motiviert in der Gruppe aktiv waren. Außerdem ein Dank an alle Studenten, die als wissenschaftliche Hilfskräfte (Katrin Müller, Melanie Busch, Jan Hellert und Sylvia Dietrich) diese Arbeit tatkräftige unterstützt haben.
Für zahlreiche Ratschläge, Diskussionen, sowie das Korrekturlesen dieser Arbeit und eine beispielhafte Zusammenarbeit danke ich meinen Laborkollegen Dr. Rebekka Biedendieck und Dr. Martin Gamer.
Ein Dank allen Labormitstreitern und Freunden für eine nicht nur fachlich konstruktive, sondern auch persönlich äußerst angenehme Umgebung während der Arbeit, aber auch bei der ein oder anderen Aktivität nach Dienstschluss.
Gedankt werden soll des Weiteren unserer Bioinformatik Gruppe (im Speziellen: Dr. Karsten Hiller, Boyke Bunk und Isam Haddad) für die Unterstützung bei den informatischen Aufgaben dieser Arbeit, sowie der Hilfe beim alltäglichen Kleinkrieg mit Soft- und Hardware.
Ein Dank allen Kooperationspartnern aus dem Sonderforschungsbereich 578 „Vom Gen zum Produkt“. Besonders danken möchte ich hier Claudia Korneli und Florian David für das Durchführen von Kultivierungen im Bioreaktor.
Ich danke der Deutsche Forschungsgemeinschaft (DFG), sowie der Firma Codexis Inc. (Redwood City, CA, USA) für die finanzielle Unterstützung dieser Arbeit.

Weiterer Dank an Prof. Dr. Rita Gerardy-Schahn und Dr. Martina Mühlenhoff von der medizinischen Hochschule Hannover, sowie Prof. Dr. Roger Hendrix und Prof. Dr. Graham Hatfull von der University of Pittsburgh für die Bereitstellung von DNA der Bakteriophagen SP6 und K1E, sowie Dr. Mark G. Carter von der University of Connecticut für das zur Verfügung stellen der pNIAysic Plasmide.
Besonderer und zugleich mein größter Dank gebührt meiner Familie, die mich sowohl persönlich als auch finanziell während meiner gesamten Ausbildungszeit stets unterstützt hat.